Cochrane
Library

Cochrane Database of Systematic Reviews

Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins (Review)

Lai NM, Foong SC, Foong WC, Tan K

Lai NM, Foong SC, Foong WC, Tan K.

Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins.

Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: CD008313.

DOI: [10.1002/14651858.CD008313.pub3](https://doi.org/10.1002/14651858.CD008313.pub3).

www.cochranelibrary.com

Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins (Review)

Copyright © 2017 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	3
BACKGROUND	6
OBJECTIVES	6
METHODS	7
RESULTS	11
Figure 1.	12
Figure 2.	14
Figure 3.	15
Figure 4.	17
Figure 5.	18
Figure 6.	19
DISCUSSION	19
AUTHORS' CONCLUSIONS	20
ACKNOWLEDGEMENTS	20
REFERENCES	21
CHARACTERISTICS OF STUDIES	24
DATA AND ANALYSES	35
Analysis 1.1. Comparison 1 Co-bedding vs standard care, Outcome 1 Rate of weight gain (in gram/kg of baseline weight/d). ...	36
Analysis 1.2. Comparison 1 Co-bedding vs standard care, Outcome 2 Apnea, bradycardia, or desaturation.	37
Analysis 1.3. Comparison 1 Co-bedding vs standard care, Outcome 3 Episodes in co-regulated state (out of 20 observations). ..	37
Analysis 1.4. Comparison 1 Co-bedding vs standard care, Outcome 4 Episodes of crying (out of 20 observations).	37
Analysis 1.5. Comparison 1 Co-bedding vs standard care, Outcome 5 Episodes in quiet sleep (out of 20 observations).	38
Analysis 1.6. Comparison 1 Co-bedding vs standard care, Outcome 6 Neurobehavior: infant pain score following painful procedure.	38
Analysis 1.7. Comparison 1 Co-bedding vs standard care, Outcome 7 Infection.	38
Analysis 1.8. Comparison 1 Co-bedding vs standard care, Outcome 8 Length of hospital stay (days).	39
Analysis 1.9. Comparison 1 Co-bedding vs standard care, Outcome 9 Parental anxiety (Parental State Anxiety Inventory).	39
Analysis 1.10. Comparison 1 Co-bedding vs standard care, Outcome 10 Parental attachment (Maternal Attachment Inventory).	40
Analysis 1.11. Comparison 1 Co-bedding vs standard care, Outcome 11 Parental satisfaction.	40
ADDITIONAL TABLES	40
APPENDICES	42
WHAT'S NEW	42
HISTORY	42
CONTRIBUTIONS OF AUTHORS	43
DECLARATIONS OF INTEREST	43
SOURCES OF SUPPORT	43
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	44
INDEX TERMS	44

[Intervention Review]

Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins

Nai Ming Lai¹, Siew Cheng Foong², Wai Cheng Foong², Kenneth Tan³¹School of Medicine, Taylor's University, Subang Jaya, Malaysia. ²Department of Paediatrics, Penang Medical College, Penang, Malaysia.³Department of Paediatrics, Monash University, Melbourne, Australia**Contact:** Nai Ming Lai, School of Medicine, Taylor's University, Subang Jaya, Malaysia. lainm@doctors.org.uk, lainm123@yahoo.co.uk.**Editorial group:** Cochrane Neonatal Group.**Publication status and date:** Edited (no change to conclusions), published in Issue 2, 2017.**Citation:** Lai NM, Foong SC, Foong WC, Tan K. Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins. *Cochrane Database of Systematic Reviews* 2016, Issue 4. Art. No.: CD008313. DOI: [10.1002/14651858.CD008313.pub3](https://doi.org/10.1002/14651858.CD008313.pub3).

Copyright © 2017 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

The increased birth rate of twins during recent decades and the improved prognosis of preterm infants have resulted in the need to explore measures that could optimize their growth and neurodevelopmental outcomes. It has been postulated that co-bedding simulates twins' intrauterine experiences in which co-regulatory behaviors between them are observed. These behaviors are proposed to benefit twins by reducing their stress, which may promote growth and development. However, in practice, uncertainty surrounds the benefit-risk profile of co-bedding.

Objectives

We aimed to assess the effectiveness of co-bedding compared with separate (individual) care for stable preterm twins in the neonatal nursery in promoting growth and neurodevelopment and reducing short- and long-term morbidities, and to determine whether co-bedding is associated with significant adverse effects.

As secondary objectives, we sought to evaluate effects of co-bedding via the following subgroup analyses: twin pairs with different weight ranges (very low birth weight [VLBW] < 1500 grams vs non-VLBW), twins with versus without significant growth discordance at birth, preterm versus borderline preterm twins, twins co-bedded in incubator versus cot at study entry, and twins randomized by twin pair versus neonatal unit.

Search methods

We used the standard search strategy of the Cochrane Neonatal Review Group (CNRG). We used keywords and medical subject headings (MeSH) to search the Cochrane Central Register of Controlled Trials (CENTRAL; 2016, Issue 2), MEDLINE (via PubMed), EMBASE (hosted by EBSCOHOST), the Cumulative Index to Nursing and Allied Health Literature (CINAHL), and references cited in our short-listed articles, up to February 29, 2016.

Selection criteria

We included randomized controlled trials with randomization by twin pair and/or by neonatal unit. We excluded cross-over studies.

Data collection and analysis

We extracted data using standard methods of the CNRG. Two review authors independently assessed the relevance and risk of bias of retrieved records. We contacted the authors of included studies to request important information missing from their published papers. We expressed our results using risk ratios (RRs) and mean differences (MDs) when appropriate, along with 95% confidence intervals (95% CIs).

We adjusted the unit of analysis from individual infants to twin pairs by averaging measurements for each twin pair (continuous outcomes) or by counting outcomes as positive if developed by either twin (dichotomous outcomes).

Main results

Six studies met the inclusion criteria; however, only five studies provided data for analysis. Four of the six included studies were small and had significant limitations in design. As each study reported outcomes differently, data for most outcomes were effectively contributed by a single study. Study authors reported no differences between co-bedded twins and twins receiving separate care in terms of rate of weight gain (MD 0.20 grams/kg/d, 95% CI -1.60 to 2.00; one study; 18 pairs of twins; evidence of low quality); apnea, bradycardia, and desaturation (A/B/D) episodes (RR 0.85, 95% CI 0.18 to 4.05; one study; 62 pairs of twins; evidence of low quality); episodes in co-regulated states (MD 0.96, 95% CI -3.44 to 5.36; one study; three pairs of twins; evidence of very low quality); suspected or proven infection (RR 0.84, 95% CI 0.30 to 2.31; three studies; 65 pairs of twins; evidence of very low quality); length of hospital stay (MD -4.90 days, 95% CI -35.23 to 25.43; one study; three pairs of twins; evidence of very low quality); and parental satisfaction measured on a scale of 0 to 55 (MD -0.38, 95% CI -4.49 to 3.73; one study; nine pairs of twins; evidence of moderate quality). Although co-bedded twins appeared to have lower pain scores 30 seconds after heel lance on a scale of 0 to 21 (MD -0.96, 95% CI -1.68 to -0.23; two studies; 117 pairs of twins; $I^2 = 75\%$; evidence of low quality), they had higher pain scores 90 seconds after the procedure (MD 1.00, 95% CI 0.14 to 1.86; one study; 62 pairs of twins). Substantial heterogeneity in the outcome of infant pain response after heel prick at 30 seconds post procedure and conflicting results at 30 and 90 seconds post procedure precluded clear conclusions.

Authors' conclusions

Evidence on the benefits and harms of co-bedding for stable preterm twins was insufficient to permit recommendations for practice. Future studies must be adequately powered to detect clinically important differences in growth and neurodevelopment. Researchers should assess harms such as infection, along with medication errors and caregiver satisfaction.

PLAIN LANGUAGE SUMMARY

Co-bedding of preterm twins to optimize growth and brain development

Review question

We reviewed evidence on the benefits and harms of co-bedding stable preterm twins and found six relevant studies.

Background

Preterm twins are at high risk for problems in growth and development. Co-bedding (placement of twins in the same cot or incubator) has been proposed to benefit twins because it simulates the environment that they shared before birth, in which twins have been seen to support each other through a series of observed activities, termed "co-regulation." These activities have been proposed to promote growth and brain development when they are allowed to continue after birth. Risks of placing twins in the same incubator or cot include caregiver error and infection.

Study characteristics

This review found six mostly small studies with some limitations in their methods. All six studies were conducted in the neonatal nursery of a major tertiary hospital. In all studies, researchers enrolled preterm infants of average postmenstrual age of 29 weeks (nearly two and a half months preterm). Some studies enrolled twins only; others enrolled twins and triplets and quadruplets and chose to co-bed two of the higher-order multiples considered most stable at the time of enrollment.

Key results

Overall, researchers reported no differences between the co-bedded group and the group receiving care separately in terms of weight gain, episodes of major disturbances in their breathing, heart rate or oxygenation level (apnea, bradycardia, or desaturation episodes), length of hospital stay, and occurrence of infection. Conflicting results were noted in the two included studies that assessed infants' pain response after heel prick.

Quality of evidence

Overall quality of evidence was low because of limitations in study methods, small sample sizes giving rise to imprecise results, and inconsistency in study results. We can make no recommendations for or against co-bedding for stable preterm twins in the neonatal nursery on the basis of evidence gathered in this review. Further research on this topic is needed.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. Co-bedding versus separate care for promoting growth and neurodevelopment in stable preterm twins

Co-bedding versus separate care for promoting growth and neurodevelopment in stable preterm twins

Patient or population: stable preterm twins with growth and neurodevelopment promoted

Settings: neonatal nursery at the hospital

Intervention: co-bedding

Comparison: separate care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Separate care	Co-bedding				
Rate of weight gain (in grams/kg of baseline weight/d) - average from study entry to week 3 Grams per kilogram per day Follow-up: 7 weeks	Mean rate of weight gain (in grams/kg of baseline weight/d) - average from study entry to week 3 in control groups was 13.7 grams per kilogram per day	Mean rate of weight gain (in grams/kg of baseline weight/d) - average from study entry to week 3 in intervention groups was 0.2 higher (1.6 lower to 2 higher)		18 (1 study)	⊕⊕⊕⊕ Low a,b	
Apnea, bradycardia, or desaturation Number of twins with episodes Follow-up: 10 days	Study population		RR 0.85 (0.18 to 4.05)	124 (1 study)	⊕⊕⊕⊕ Low b,c	
	53 per 1000	45 per 1000 (9 to 213)				
	Moderate					
	53 per 1000	45 per 1000 (10 to 215)				
Episodes in co-regulated state (out of 20 observations) Episodes observed (out of a total of 20 observations) Scale from 0 to 20 Follow-up: 2 months	Mean number of episodes in co-regulated state (out of 20 observations) in control groups was 13.94 episodes	Mean number of episodes in co-regulated state (out of 20 observations) in intervention groups was 0.96 higher (3.44 lower to 5.36 higher)		6 (1 study)	⊕⊕⊕⊕ Very low b,d	

Infections - suspected or proven infections (any) Clinical and microbiological methods Follow-up: 2 months	Study population		RR 0.84 (0.3 to 2.31)	65 (3 studies)	⊕⊕⊕⊕ Very low a,b,d
	182 per 1000	153 per 1000 (55 to 420)			
	Moderate				
	100 per 1000	84 per 1000 (30 to 231)			
Length of hospital stay (days) Follow-up: mean 2 months	Mean length of hospital stay (days) in control groups was 52.7 days	Mean length of hospital stay (days) in intervention groups was 4.9 lower (35.23 lower to 25.43 higher)		6 (1 study)	⊕⊕⊕⊕ Very low a,b,d
Parental satisfaction Parental satisfaction survey Scale from 0 to 55 Follow-up: 5 days	Mean parental satisfaction in control groups was 51.13 points (out of 55)	Mean parental satisfaction in intervention groups was 0.38 lower (4.49 lower to 3.73 higher)		18 (1 study)	⊕⊕⊕⊕ Moderate e
Neurobehavior: infant pain score following painful procedure - at 30 seconds post heel lance Premature Infant Pain Profile (PIPP) Scale from 0 to 21	Mean neurobehavior: infant pain score after painful procedure - at 30 seconds post heel lance in control groups was PIPP scale (0 to 21)	Mean neurobehavior: infant pain score after painful procedure - at 30 seconds post heel lance in intervention groups was 0.96 lower (1.68 to 0.23 lower)		224 (2 studies)	⊕⊕⊕⊕ Low f,g

*The basis for the **assumed risk** (e.g. median control group risk across studies) is provided in footnotes. The **corresponding risk** (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)

CI: confidence interval; **RR:** risk ratio

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate

Very low quality: We are very uncertain about the estimate

^aDowngraded 1 level for study limitations (owing to unclear method of sequence generation, non-blinding, and presence of selective outcome reporting)

^bDowngraded 1 level for imprecision (owing to wide 95% confidence interval, which encompassed effect sizes ranging from a clear and/or up to a moderately large benefit for the co-bedded group to a clear and/or up to a moderately large benefit for the control group)

- ^cDowngraded 1 level for indirectness (as the study assessed specifically A/B/D episodes after a painful procedure in the form of a heel lance)
- ^dDowngraded 2 levels for serious study limitations (as this was a pilot study that was non-blinded, with high risk of bias for incomplete outcome data and selective outcome reporting)
- ^eDowngraded 1 level for study limitations (owing to unclear sequence generation and allocation concealment and non-blinding)
- ^fDowngraded 1 level for study limitations (owing to non-blinding of care personnel and outcome assessors)
- ^gDowngraded 1 level for inconsistency (owing to differing direction of results between the 2 included studies, with $I^2 = 75\%$)

BACKGROUND

Description of the condition

Advances in antenatal care and use of assisted fertilization techniques (Blondel 2002) over the past three decades have led to an increase in the rate of birth of twins and higher-order multiples (Imaizumi 1998; Hur 2005; Eriksson 2007), many of whom are preterm infants (Buscher 2000; Blondel 2002a). Compared with singletons of the same postmenstrual age (PMA), twins are at higher risk for intrauterine growth restriction (Liu 2002; Jacquemyn 2003) and developmental disabilities (Topp 2004; Bonellie 2005; Glinianaia 2006). Advances in neonatal intensive care and nutrition have played a major role in improving the prognosis of these infants; however, other measures that may optimize their growth and neurodevelopment should be rigorously assessed.

It is postulated that twins interact with each other while in utero via a series of activities collectively termed "co-regulation," conferring mutual support that might benefit their growth and development (Nyqvist 1998). Co-regulatory activities observed include touching, holding, rooting, and hugging, among others (Nyqvist 1998). Co-regulation might be due in part to the physical proximity of these fetuses in utero, which is lost when infants are nursed separately after birth. For preterm twins, the stress of coping with the extrauterine environment alone may compound the problems faced by preterm infants, leading to even greater deficits in growth and neurodevelopment (Hayward 2003). Some evidence suggests that measures taken to reduce stress for a newborn may improve neurodevelopment and parental satisfaction (Feldman 2002).

Description of the intervention

Co-bedding is defined as care of two or more hospitalized infants in the same cot or incubator. Co-bedded infants may be nursed clothed or naked, and periodic care procedures, such as feeding, cleaning, and changing, may be performed simultaneously on all infants by multiple nurses or sequentially by the same nurse, as dictated by the practice protocol of the unit. Nursing environments when co-bedding is initiated, including incubator temperature and humidity and ongoing monitoring of co-bedded infants, should be the same as when infants are cared for separately. Co-bedding is currently practiced only with stable infants because of the difficulties and risks associated with caring for two sick infants in the same incubator. Co-bedding may have to be suspended when any or all of the co-bedded infants develop a medical complication, such as infection or the need for phototherapy. When the problem is resolved, co-bedding may be re-instituted.

How the intervention might work

Co-bedding of twins has been introduced in an attempt to maintain their in utero proximity, allowing continuation of fetal co-regulation, thus easing the transition from intrauterine to extrauterine life. Based on the belief that twins may have a special capacity for supporting each other because of their common intrauterine experiences, this series of activities is believed to promote growth and neurodevelopment (Nyqvist 1998; Hayward 2003), which are chief measures of well-being for stable infants. Exactly how these activities lead to improved growth and neurodevelopment is unknown.

Other proposed advantages of co-bedding include heat conservation, apnea reduction, improved bonding between

infants, improved parental and staff satisfaction (Nyqvist 1998), and saving of cot space. Preterm twins are likely to be the predominant participants in studies on co-bedding because of their relative prevalence and cot capacity and longer hospital stay compared with full-term twins. Among them, twins with significant growth discordance (> 30% difference in birth weight) represent a subgroup of interest for this intervention, as they have been shown consistently to have greater mortality and increased growth and developmental morbidities at various stages of life compared with twins with less growth discordance (Cheung 1995; Scher 2002; Goyen 2003; Adegbite 2004; Yinon 2005).

Why it is important to do this review

Traditionally practiced in many European countries, co-bedding has been seen increasingly in American neonatal units since the 1990s (Hayward 2003; Tomashek 2007). Reports of co-bedding for hospitalized neonates in other areas of the world have been scarce.

Concerns on co-bedding include risk of infection, possible medication error, difficulties involving the routine care of individual infants, less reliable incubator temperature control, mutual disturbance leading to reduced sleep quality, and possible physical injury (Hayward 2003).

The magnitude of the benefits and risks of co-bedding is currently unclear, and benefits and risks should be evaluated in a variety of clinical situations. Subgroup analyses should be undertaken to evaluate whether effects of co-bedding are different between groups of infants with different key prognostic characteristics, for example, birth weight (very low birth weight [VLBW] vs non-VLBW) and postmenstrual age (PMA) (Gutbrod 2000; Hack 2003), growth discordance (Bagchi 2006), and nursed in incubator or cot (Gray 2003), and between twins in a unit where co-bedding is a standard practice as opposed to a unit where co-bedding and separate care are practiced in the same setting.

Although infant sleep habits are difficult to monitor beyond the time of discharge, long-term effects of hospital intervention, even for short periods, cannot be excluded. Co-bedding of twins, if practiced in the hospital nursery, may influence parental practices at home. This necessitates assessment of long-term growth, neurodevelopment, and infant mortality, especially in view of general concerns surrounding reports linking infant sleeping practices as well as twinning with sudden infant death syndrome (SIDS) (American Academy of Pediatrics 2005; Gilbert 2005; Adler 2006; Pharoah 2007).

Co-bedding is considered part of an infant's developmental care (Byers 2003b), especially for a specific and growing target population, namely, twins. Symington and colleagues (Symington 2006) have reviewed various aspects of individualized developmental care of the neonate, but continued research is needed to assess the benefits and risks of co-bedding for stable preterm twins in terms of growth, neurodevelopment, and other clinically relevant outcomes. If found beneficial and safe, co-bedding will prove to be a valuable recommendation, particularly for neonatal units with a shortage of infant cot space.

OBJECTIVES

We aimed to assess the effectiveness of co-bedding compared with separate (individual) care for stable preterm twins in the neonatal

nursery in promoting growth and neurodevelopment and reducing short- and long-term morbidities, and to determine whether co-bedding is associated with significant adverse effects.

As secondary objectives, we sought to evaluate effects of co-bedding via the following subgroup analyses: twin pairs with different weight ranges (very low birth weight [VLBW] < 1500 grams vs non-VLBW), twins with versus without significant growth discordance at birth, preterm versus borderline preterm twins, twins co-bedded in incubator versus cot at study entry, and twins randomized by twin pair versus neonatal unit.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomized controlled trials and quasi-randomized trials at the level of each twin pair or at the level of each neonatal unit (neonatal intensive care unit [NICU]-cluster trials). We excluded cross-over trials, as we considered it highly likely that the major outcomes examined, such as weight gain and neurobehavior, would be significantly "contaminated" by the cross-over.

Types of participants

We included studies that enrolled pairs of twins of less than 37 weeks' gestation who were considered clinically stable at study entry. We excluded studies that exclusively enrolled infants under intensive care (including those needing any type of respiratory support apart from nasal prong, low-flow oxygen supplement, and/or any type of circulatory support) or those clinically considered septic at study entry or with congenital anomaly or open skin wounds. For studies with a mixture of stable and less stable infants, we planned to retrieve and analyze only data from stable infants; we planned to contact the primary study authors if necessary.

We placed no limit on weight and did not restrict the lower limit of PMA.

We included both growth-concordant and growth-discordant twins in the overall analysis and separated significantly growth-discordant twins from others in subgroup analyses. We defined significant growth discordance in this review as an intra-pair difference in birth weight greater than 30%, with the weight of the larger twin taken as the denominator, as weight discordance beyond this level has been consistently associated with worsening of clinically important outcomes (Cheung 1995; Yalcin 1998; Branum 2003; Bagchi 2006).

Participants were eligible regardless of their postnatal age as soon as they were considered clinically stable by the attending physicians.

Types of interventions

Intervention

Co-bedding was defined as placement of twins in the same cot or incubator for nursery care. Co-bedded twins might be nursed clothed or naked, and periodic care procedures, such as feeding, cleaning, and changing, might be performed simultaneously on both twins by two nurses or sequentially by the same nurse, as dictated by the practice protocol of the unit. If any twin required separate care as the result of a problem such as infection or the

need for phototherapy, co-bedding might be re-instituted when the problem was resolved, as long as the twins were co-bedded for a total duration of at least one week, as specified under the subheading "Length of intervention."

Time of initiation

We included twin pairs of any postnatal age if they were considered clinically stable as defined by study authors.

Length of intervention

We did not set a minimum duration of intervention because it is unknown how long twins need to be in close contact with each other to show meaningful effects on physiological parameters, growth, and neurodevelopmental outcomes, and studies that have assessed skin-to-skin contact between newborn infants and their mothers show that beneficial effects can be demonstrated after as little as 15 minutes of such contact (Moore 2012). We accepted any timing for commencement of the intervention, as long as infants were considered clinically stable when the intervention was started. Once the intervention was instituted, twins had to be co-bedded the whole time, except during nursing care or contact with family members, such as during breast-feeding and parental cuddling.

Comparison

Each twin was placed in a separate cot or incubator for nursery care.

Adjustments to the nursing environment involving incubator temperature and humidity and ongoing monitoring should be guided by the same protocol for co-bedded twins as for twins nursed separately on the same unit. The sleeping position of each infant participant (supine, prone, or lateral), regardless of assigned group, was standardized.

We excluded trials that examined co-bedding with concurrent developmental care interventions, such as nesting, swaddling, kangaroo care, and the Neonatal Individualized Developmental Care Program (NIDCAP) regimen (Symington 2006), because use of these measures likely limits the freedom of co-bedded twins in undertaking co-regulatory activities, thereby weakening the credibility of the claim of association between any observed effect and mechanisms through which co-bedding was postulated to work.

Types of outcome measures

We have listed units of analysis for the following outcomes for studies in which randomization occurred at the level of each twin pair. We have provided additional details on how we handled units of analysis under the heading [Unit of analysis issues](#).

Primary outcomes

- Growth: average rate of weight gain between twin pairs for the whole period of the study, expressed in grams/kg/d. We used a standard formula to derive the rate of weight gain (Patel 2005).
- Apnea: average number of episodes between twin pairs. Apnea was defined as cessation of breathing for at least 20 seconds, or for a shorter period if accompanied by bradycardia or cyanosis (American Academy of Pediatrics 2003).
- Bradycardia: average number of episodes between twin pairs. Bradycardia was defined as a fall in heart rate of greater than

- 30% to below baseline ([Henderson-Smart 1986](#)) or to less than 100 beats per minute for at least 10 seconds ([Bredemeyer 2012](#)).
- Neurobehavior in the nursery measured by validated scales such as the Brazelton Neonatal Behavior Assessment Scale (NBAS) and the Neonatal Intensive Care Unit Network Neurobehavioral Scale (NNNS) ([Lester 2004](#)), including infant co-regulation, stress response, and sleep/wake state, as defined by the proportion of time in quiet sleep.

Secondary outcomes

- *Number of twin pairs with sepsis (clinically diagnosed and/or microbiologically proven).
- Average number of days between twin pairs to reach full enteral feed by gavage or by mouth at a volume of at least 150 mL/kg/d.
- *Number of twin pairs with feed intolerance (defined as a need to stop enteral feeds).
- *Number of twin pairs requiring intensive care, including the need for any type of respiratory and/or circulatory support.
- *Number of twin pairs for whom cessation of co-bedding is required for other clinical reasons, such as infection or the need for phototherapy.
- All-cause neonatal mortality (mortality in the period up to 28 days after birth) in one or both twins, before or after discharge.
- Average length of hospital stay between twin pairs (days).
- Average cost of care between twin pairs.
- Long-term growth, as indicated by average weight gain between twin pairs measured at intervals as specified by the authors of individual studies, for example, at three, six, 12 and 18 months.
- Long-term neurodevelopment, measured by validated scales such as Bayley Scales of Infant Development ([Washington 1998](#)), whereby average scores between twin pairs would be taken if data were available. Clinically diagnosed non-ambulatory cerebral palsy or significant auditory and visual impairment would be accepted if data were available.
- All-cause infant mortality in one or both twins.
- SIDS in one or both twins.
- Parental and staff satisfaction measured by validated scales, such as a health-related quality of life tool (HRQoL) or a health outcome rating scale, or self reported parental or staff dissatisfaction. One overall measure would be taken for the parents of each twin pair if data were available.
- Adverse effects (number of twin pairs in whom these occur), such as medication errors, hypothermia (core body temperature < 36.5°C or skin temperature < 36°C), or hyperthermia (core temperature > 38°C or skin temperature > 37.5°C) ([WHO 1997](#)), physical injury, and other patient safety-related outcomes as defined by individual studies.
- Incidence of co-bedding post discharge.
- Other clinically relevant outcomes, as specified by the authors of individual studies.

*For these secondary outcomes, we would count as one positive outcome incidents when one or both twins developed the stated outcome.

Outcomes were measured at the following points in time.

- During hospital stay (weight gain, neurobehavioral status, clinically important outcomes such as sepsis and feed intolerance, parental and staff satisfaction).
- At discharge (weight, parental and staff satisfaction).
- At defined points beyond discharge, as variously set by the authors of included studies when assessment of growth and neurodevelopment becomes feasible.

Search methods for identification of studies

We used the criteria and standard methods of The Cochrane Collaboration and the Cochrane Neonatal Review Group (see the [Cochrane Neonatal search strategy for specialized register](#)).

Electronic searches

We conducted a comprehensive search of the Cochrane Central Register of Controlled Trials (CENTRAL; 2016, Issue 2); MEDLINE via PubMed (1966 to February 29, 2016); EMBASE (1980 to February 29, 2016); and the Cumulative Index to Nursing and Allied Health Literature (CINAHL) (1982 to February 29, 2016) using the following search terms: (Nurseries OR Neonatal nursing OR NICU) AND (cobed* OR (co bed*) OR (bed shar*) OR (co sleep*)), plus database-specific limiters for randomized controlled trials (RCTs) and neonates (see [Appendix 1](#) for the full search strategy for each database). We applied no language restrictions.

We searched clinical trials registries for ongoing or recently completed trials ([clinicaltrials.gov](#), the World Health Organization International Trials Registry and Platform [[www.who.int/ictrp/search/en/](#)], and the International Standard Registered Clinical Trial Number [ISRCTN] Registry [[ISRCTN Registry](#)]).

We accepted studies whether published or unpublished, in full article or abstract form, as long as assessment of study quality was possible and other inclusion criteria had been fulfilled.

Searching other resources

We contacted the authors of all selected studies about ongoing trials and unpublished data. We searched the references cited in previous relevant Cochrane reviews and in other relevant papers, review articles, and Internet documents. We assessed the results of handsearching by the Cochrane Neonatal Review Group. We searched the proceedings of relevant pediatric and perinatal conferences (specifically, those of the American Academy of Pediatrics, the European Academy of Pediatrics, the European Society of Paediatric Research (ESPR), the United Kingdom Perinatal Society, and the Perinatal Society of Australia and New Zealand) to identify relevant published abstracts.

Data collection and analysis

We followed the standards of the Cochrane Neonatal Review Group.

Quality of evidence

We assessed the quality of evidence for the main comparison at the outcome level using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach ([Guyatt 2011a](#)). This methodological approach considers evidence from randomized controlled trials to be of high quality that may be downgraded on the basis of consideration of any of five areas: design (risk of bias), consistency across studies, directness of evidence, precision of estimates, and presence of publication bias

(Guyatt 2011a). The GRADE approach assigns the quality of a body of evidence to one of the following grades: High quality: We are very confident that the true effect lies close to that of the estimate of the effect; Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but it may be substantially different; Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect; and Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect (Schünemann 2013).

Review authors independently assessed the quality of the evidence for outcomes identified as critical or important for clinical decision making. These outcomes include rate of weight gain (in gram/kg of baseline weight/d), average from study entry to week three; apnea, bradycardia, or desaturation; episodes in the co-regulated state (out of 20 observations); infections (any suspected or proven infection); length of hospital stay (in days); parental satisfaction; and infant pain score following a painful procedure (at 30 seconds post heel lance).

In cases in which we considered that the risk of bias resulted from inadequacies in concealment of allocation, randomized assignment, complete follow-up, or blinded outcome assessment, our confidence in effect estimates was reduced, and we downgraded the quality of evidence accordingly (Guyatt 2011b). We evaluated consistency by noting similarity of point estimates, extent of overlap of confidence intervals, and statistical criteria, such as measurement of heterogeneity (I^2). We downgraded the quality of evidence when large and unexplained inconsistency was observed across studies (i.e. some studies suggest important benefit, and others no effect or harm without a clinical explanation) (Guyatt 2011d). We assessed precision on the basis of width of the 95% confidence interval (CI) and the calculated optimal information size (OIS). If the total number of participants included in the pooled effect estimation was less than the number of participants generated by a conventional sample size calculation for a single adequately powered trial, we considered rating the quality of evidence down for imprecision (Guyatt 2011c). When trials were conducted in populations other than the target population, we downgraded the quality of evidence because of indirectness (Guyatt 2011e).

We entered data (i.e. pooled estimates of effects and corresponding 95% CIs) and explicit judgments for each of the above assessed aspects into the Guideline Development Tool - the software used to create "Summary of findings" tables (GRADEpro). We explained all judgments involving assessment of the study characteristics described above in footnotes or comments provided in the "Summary of findings" tables.

Selection of studies

We used standard methods of The Cochrane Collaboration and its Neonatal Review Group. All review authors independently assessed studies and determined their eligibility by using a form with predefined inclusion criteria. We excluded any study that did not fit the criteria and resolved disagreements by discussion.

Data extraction and management

Three review authors (NML, SCF, and WCF) independently entered individual data from each included trial onto a standardized data collection form. Data entered were compared, and differences were resolved by consensus.

Assessment of risk of bias in included studies

Two review authors (NML and SCF) independently assessed risks of bias of included trials, using standard methods of The Cochrane Collaboration as presented in the *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1 (Higgins 2011).

Specifically, we assessed trials on the basis of the following: adequacy of sequence generation, allocation concealment, blinding to the intervention, completeness of follow-up, and blinding to outcome measurement.

Had data been available from cluster-randomized trials, we planned to assess the following specific additional biases when applicable: recruitment bias, baseline imbalance, loss of clusters or individuals within clusters, incorrect unit of analysis (as elaborated in the corresponding section), and compatibility with individually randomized trials.

Information on risks of bias was included in the [Characteristics of included studies](#) table.

In addition, we evaluated the following issues and made a judgment regarding high, unclear, or low risk of bias.

- Sequence generation: Was the allocation sequence adequately generated?
- Allocation concealment: Was allocation adequately concealed?
- Blinding of participants, personnel, and outcome assessors: Was knowledge of the allocated intervention adequately prevented during the study? At study entry? At the time of outcome assessment?
- Incomplete outcome data: Were incomplete outcome data adequately addressed?
- Selective outcome reporting: Were reports of the study free of the suggestion of selective outcome reporting?
- Other sources of bias: Was the study apparently free of other problems that could put it at high risk of bias?

Measures of treatment effect

We followed the procedures of the Cochrane Neonatal Review Group. For categorical data, when applicable, we used risk ratios (RRs), risk differences (RDs), and numbers needed to treat for an additional beneficial outcome (NNTBs) along with their respective 95% confidence intervals (CIs). For continuous data, such as length of hospital stay, we used mean differences (MDs) with 95% CIs.

Unit of analysis issues

We identified the unit of analysis for each individual trial to determine whether appropriate adjustments had been made to account for non-independence between twin pairs via use of appropriate analysis methods, such as generalized estimating equation (GEE) modeling. Had an NICU-cluster trial been included, we would have looked for evidence of adjustments at the level of the NICU, in addition to adjustments at the level of each twin

pair. If the unit of analysis was not stated in the study, we planned to inspect the width of the standard error (SE) or the 95% CI of estimated treatment effects. If we found an inappropriately small SE or a narrow 95% CI, we would have asked the authors of the study to provide information on the unit of analysis.

In this review, we considered each twin pair as representing one unit when we analyzed data from trials in which randomization occurred at the level of each twin pair. For continuous outcomes, we reported the average figure for each twin pair, for example, the average rate of weight gain of the twin pairs. For dichotomous outcomes with low event rates, we counted the event as positive if either of the twins developed the outcome (e.g. in this review, we counted the outcome of sepsis as positive if one or both twins developed sepsis). For dichotomous outcomes with high event rates, had data been available, we would have assigned a score of zero, one, or two, which would correspond to the number of twins in a twin pair with the outcome of interest.

When we found a unit of analysis error, for example, when the included study analyzed outcome data for individual infants instead of for each twin pair, we re-analysed data derived from the studies and adjusted for the effect of clustering by using different strategies for two main types of studies, as follows.

- For studies in which randomization occurred at the level of each twin pair, as in the included studies, we adjusted individual data, counting each twin dyad as an individual. For continuous data, we recorded the average reading of the twin pair. For dichotomous data, we recorded the number of twin pairs with the outcome (positive if any or both of the twins developed the outcome).
- Had we included studies in which randomization occurred at the level of the neonatal unit (i.e. NICU-cluster trials), we would have dealt with two levels of clustering: (a) clustering at the level of the twin pair, and (b) clustering at the level of the neonatal unit.
 - To address (a), we would have adjusted the individual data, counting each twin dyad as an individual. For continuous data, we would have recorded the average reading of the twin pair. For dichotomous data, we would have recorded the number of twin pairs with the outcome (positive if any of the twins developed the outcome).
 - To address (b), we would have adjusted the final estimates of each study obtained by analyzing each twin dyad, as stated. We would have done so by multiplying the SEs of the final effect estimates by the square root of the "design effect": $(1 + (M - 1) \times ICC)$, where M was the average cluster size (number of twin pairs per cluster) and ICC was the intra-cluster correlation among twin pairs within each NICU. We would have determined the average cluster size (M) from each trial by dividing the total number of twin pairs by the total number of clusters. We would have sought the best estimate of the ICC from reliable resources such as landmark studies involving twins and NICU clustering.
 - We would have combined adjusted final effect estimates from each trial with their SEs in the meta-analysis using generic inverse variance methods.

Should we fail to identify a reliable ICC for NICU-cluster trials for twins after searching through relevant resources, we planned to use partially adjusted estimates (adjusted only at the level of each twin dyad) for our meta-analysis. We planned to include studies

for which determination of the unit of analysis was not possible in the overall meta-analysis on the basis of effect estimates given by study authors. We would have then performed a sensitivity analysis to assess how the overall results were affected with and without inclusion of these studies.

Dealing with missing data

We obtained a drop-out rate from each included study. A drop-out rate equal to or greater than the control group event rate would be considered significant. If a significant drop-out rate had been found, we would have contacted the authors of the individual studies to ask that they provide additional data when possible.

Had a sufficient quantity of data been available, we would have performed a sensitivity analysis to assess how the overall results were affected with and without inclusion of studies with significant drop-out rates.

Assessment of heterogeneity

Had we included more than one study that reported the same outcome measure, we would have assessed the treatment effects of individual trials and heterogeneity between trial results by first inspecting the forest plots.

We then would have evaluated the I^2 statistics to look for an indication of the degree of heterogeneity in study results, with a value greater than 50% indicating moderate or high heterogeneity. If moderate or high heterogeneity was found, we would have explored the causes (e.g. differences in study quality, participants, interventions, or outcome assessments) via post hoc subgroup analyses. We then would have made a decision about whether a meta-analysis would be appropriate and, if so, whether we should exclude some studies.

Assessment of reporting biases

We constructed a matrix consisting of all outcomes assessed in each included study, as stated in the methods and reported in the results. We compared reported outcomes against outcomes listed in the methods section. In addition, when we found in a study some key outcomes that we considered assessable but were not included, we contacted the study authors to ask for clarification. We sought the study protocol, when possible, from PubMed, a relevant trial registry, or a Web link provided by the study, or we directly contacted study authors. We performed cross-examination against the conference presentation of the same work, when possible. We planned that if all of the above measures failed to provide further useful information, we would perform a sensitivity analysis by taking an outcome that was reported by all studies and comparing overall results with and without inclusion of studies in which key outcomes were missing.

We planned to screen for publication bias by using a funnel plot if we included a sufficient number of studies (≥ 10) in the analysis. If we suspected publication bias (i.e. significant asymmetry upon visual inspection of the funnel plot), we would have included a statement in our results and a corresponding note of caution in our discussion.

Data synthesis

Had data from more than one study been available for each outcome, we would have used a fixed-effect model to combine the

final effect estimates from each trial with their SEs in the meta-analysis using generic inverse variance methods in RevMan.

Subgroup analysis and investigation of heterogeneity

Had relevant data been available, we would have performed subgroup analyses for the following.

- Twins with different weight range: VLBW (birth weight of the smaller twin < 1500 g) vs non-VLBW.
- Twins with and without significant weight discordance at birth.
- Preterm twins (< 34 completed weeks of PMA at commencement of co-bedding) and twins of borderline maturity (35 and 36 weeks' completed PMA).
- Twins co-bedded in an incubator and a cot at study entry.
- Studies in which randomization occurred at the level of each twin pair and each neonatal unit.

Sensitivity analysis

We would have performed sensitivity analyses based on the following, if applicable.

- Adequacy of sequence generation and allocation concealment, as detailed under the heading [Assessment of risk of bias in included studies](#).
- Adequacy in handling missing data, as detailed under the heading [Dealing with missing data](#).
- Availability and adequacy of adjustments of units of analysis, as detailed under the heading [Unit of analysis issues](#).

RESULTS

Description of studies

Results of the search

In the original search, we identified a total of 344 records, including 85 records from PubMed, 80 from EMBASE, 14 from CENTRAL, and 165 from CINAHL. A search of CINAHL based on the equivalent search strategy yielded only eight records. Without publication type

filters (item #11 in our search strategy [see [Electronic searches](#)]), the CINAHL search yielded 165 records. We decided to screen those 165 records alongside records identified through the other databases. We found two additional relevant articles by reviewing the reference list of a review article ([Tomashek 2007](#)) and noted that both publications ([Lutes 2000a](#); [Longobucco 2002](#)) were not indexed in any of the databases that we had searched. After removal of duplicates, we retained 302 records for further inspection. For our 2016 update, we identified 47 additional records, including 41 from the databases and six from the trial registries. After removal of duplicates, 34 records were available for further inspection.

After inspecting the titles, we short-listed 25 articles, including four from our 2016 update. These included 14 articles from PubMed, six from CINAHL, three from EMBASE, and both additional references as stated above ([Lutes 2000a](#); [Longobucco 2002](#)). We found no additional relevant records in CENTRAL. One article that we identified through PubMed included a protocol of a study, which was subsequently published in abstract form and in full text ([Campbell-Yeo 2012b](#)), along with three other reports of the same study with outcomes that were not pre-specified as relevant in our review; these were included as additional references for [Campbell-Yeo 2012b](#). We also identified an article that was published in French ([Mazeiras 2010](#)) - the only non-English article in our short list.

Two articles in our short list ([Lutes 2000](#); [Lutes 2001](#)) appeared to describe the same study, but we could not confirm this, as we were unable to acquire the full text of [Lutes 2000](#), and we have not heard from the study authors despite multiple emails sent since the initial review was drafted to request further information. We have excluded [Lutes 2000](#) from our 2016 update for lack of information.

After we had assessed the abstracts or full-text writings of these short-listed articles, we included five articles ([Lutes 2001](#); [Byers 2003a](#); [Chin 2006](#); [Hayward 2007](#); [Campbell-Yeo 2012b](#)) in our original analysis along with one additional article ([Badiee 2014](#)) from our 2016 update. We have provided in [Figure 1](#) a flow diagram of the review process from initial search to final inclusion of studies.

Figure 1. Flow diagram of the review process from initial search to final inclusion of studies.

Figure 1. (Continued)

Included studies

All six included studies were randomized controlled trials conducted in level three NICUs, with co-bedding instituted according to standard protocols. [Byers 2003a](#) enrolled eight pairs of co-bedded infants and 21 infants cared for by separate nursing, as well as 19 parents. All infants were preterm (average 29 weeks' PMA). A mixture of twins and triplets were enrolled, and two triplets who were considered most stable at the time of enrollment were chosen for co-bedding. [Campbell-Yeo 2012b](#) enrolled 67 sets of twins at between 28 and 36 weeks' PMA. In [Chin 2006](#), participants consisted of 39 pairs of twins at less than 34 weeks' PMA. [Hayward 2007](#) was a pilot study in which investigators assessed six pairs of twins weighing less than 2000 grams. [Lutes 2001](#) recruited 38 sets of twins, 11 sets of triplets and three sets of quadruplets. [Badiie 2014](#) recruited 50 pairs of twins of 26 to 34 weeks' PMA whose postnatal age was less than 20 days.

In [Byers 2003a](#), major outcomes were infants' physiological parameters, sleep-wake synchrony, and parental perception of care. [Campbell-Yeo 2012b](#) and [Badiie 2014](#) assessed specifically the response of infants to heel lance by using the Premature Infant Pain Profile (PIPP), which could be classified under infant neurobehavior. [Chin 2006](#) included weight and combined episodes of apnea/bradycardia and desaturation (A/B/D) as the primary outcomes. Infection and medication error were the secondary outcomes in this study. [Hayward 2007](#) reported amount of quiet sleep, crying, and time spent co-regulated as major outcomes. Study authors employed a rating scale (0 to 20) for these outcomes that was based on the number of episodes in which outcomes of interest were present during the designated 20 periods of observation (see [Characteristics of included studies: Hayward](#)

[2007](#), "Outcomes," for a more detailed description of the unit of measurement). [Lutes 2001](#) reported weight gain, growth of head circumference, and length as primary outcomes. Secondary outcomes were medication error, nosocomial infection, sepsis workup initiated, and thermal insult. However, study authors did not report outcomes separately for twins, triplets, and quadruplets. We have provided a detailed description of all studies in the [Characteristics of included studies](#) table.

In all six studies, the unit of analysis was the individual infant rather than the twin pair, and two studies ([Campbell-Yeo 2012b](#); [Badiie 2014](#)) performed statistical adjustments to account for the effects of twin clustering by using GEE. We adjusted the unit of analysis for [Chin 2006](#) after re-analysing the raw data that we obtained from the lead study author. We accounted for the unit of analysis in [Hayward 2007](#) by adjusting the number of participants from the total number of individual infants to the total number of twin pairs. We are still waiting for the authors of [Byers 2003a](#) and [Lutes 2001](#) to provide us with their raw data so we can make adjustments for the unit of analysis.

Excluded studies

Among the 15 excluded studies, four ([DellaPorta 1998](#); [Anon 2002](#); [Wahl 2006](#); [Mazeiras 2010](#)) were commentaries or review articles, four ([Fischer 1991](#); [Nyqvist 2002](#); [Matthews 2004](#); [Jahanfar 2012](#)) were original studies that did not address our research question on co-bedding of preterm infants, and six assessed co-bedding of preterm infants, but all were prospective single-group studies ([Nyqvist 1998](#); [Touch 2002](#)), a prospective cross-over study ([Stainton 2005](#)), a prospective study with historical controls ([Longobucco 2002](#)), or retrospective cohort studies ([Polizzi 2003](#);

LaMar 2006). We excluded one study (Lutes 2000) from our 2014 update when we were unable to obtain further information; we could not obtain the full text nor a reply from study authors after multiple attempts. We have provided a detailed description of these studies in the [Characteristics of excluded studies](#) table.

Risk of bias in included studies

We have provided a graphical summary of the overall risk of bias in [Figure 2](#) and [Figure 3](#).

Figure 2. Methodological quality summary: review authors' judgements about each methodological quality item for each included study.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)
Badiee 2014	+	+	+	+	+	-	-
Byers 2003a	?	?	+	+	-	-	?
Campbell-Yeo 2012b	+	+	+	+	+	-	?
Chin 2006	?	+	+	-	+	-	-
Hayward 2007	+	+	-	-	-	-	-
Lutes 2001	?	?	-	-	+	-	?

Figure 3. Methodological quality graph: review authors' judgements about each methodological quality item presented as percentages across all included studies.

Allocation

Two articles ([Campbell-Yeo 2012b](#); [Badiie 2014](#)) provided sufficient details of random sequence generation and allocation. [Chin 2006](#) achieved randomisation by using randomly permuted blocks of two, four, and six via a computerized off-site Website accessed only by the principal investigator or a research nurse. [Badiie 2014](#) performed randomization by using computer-generated random number tables and carried out allocation by using sealed and opaque envelopes. The authors of both [Chin 2006](#) and [Hayward 2007](#) provided further details in correspondence describing how allocations were carried out. [Chin 2006](#) used sealed and opaque envelopes for allocation, and [Hayward 2007](#) placed the assignment in folded cards that were taped shut and placed into sealed envelopes. [Byers 2003a](#) and [Lutes 2001](#) provided no details on how randomization was achieved. Overall, two studies had low risk and three studies had unclear risk of bias on sequence generation, and three studies had low risk of bias on allocation concealment.

[Byers 2003a](#), [Chin 2006](#), and [Hayward 2007](#) provided evidence indicating that allocated groups were unequal in some major characteristics. In [Byers 2003a](#), co-bedded infants had higher PMA at study enrollment compared with infants in the control group (average 33 weeks' PMA in co-bedding group and 30.6 weeks' PMA in control group) and higher proportions of male infants (68.3% in the co-bedded group and 22.7% in the control group). In [Campbell-Yeo 2012b](#), participants appeared well matched in their baseline characteristics. The authors of [Chin 2006](#) stated that "the two groups were well balanced for PMA, gender, and race, although not for baseline weight" (Results, paragraph 4, lines 1 to 3), and in [Hayward 2007](#), study authors stated, "There was a difference in the distribution by age, education, and employment status" (Sample characteristics, paragraph 1, lines 4 to 6). [Lutes 2001](#) did not comment on the degree of similarity in baseline characteristics of their sample. Table 4 in the article shows that the two groups appeared well balanced in baseline characteristics, except in mean birth weight (co-bedded group 1459.3 grams, control group 1262.6 grams).

Blinding

Although it was not clearly stated in most studies, blinding of caregivers as well as outcome assessors appeared very unlikely across all studies in view of the nature of the intervention assessed. Further correspondence with the lead authors of [Chin 2006](#) and [Hayward 2007](#) revealed that data collectors were not blinded to the allocation status of infants. We are still awaiting response from the authors of [Byers 2003a](#) and [Lutes 2001](#) to clarify whether data collectors were blinded.

Incomplete outcome data

Four studies ([Byers 2003a](#); [Chin 2006](#); [Campbell-Yeo 2012b](#); [Badiie 2014](#)) had low risk of bias, and two studies ([Lutes 2001](#); [Hayward 2007](#)) had high risk of bias. In [Badiie 2014](#), 100 out of 105 recruited infants (95.2%) completed the study, although the five infants who did not receive the allocated intervention were dropped from the analysis with no reasons stated. In [Byers 2003a](#), all except one infant, who was diagnosed with sepsis, completed the study. [Campbell-Yeo 2012b](#) reported a 7.4% non-completion rate and a balanced number of non-completers between the two groups. [Chin 2006](#) started with 41 pairs of twins and ended with 22 pairs of twins at last measurements. Study authors accounted for the progressive loss of data by stating that some infants had been discharged home over the study period. [Hayward 2007](#) started with six pairs of twins and ended with two pairs of twins at last measurements. The study author stated upon further contact that loss of data was due to early discharge and unplanned transfer to other hospitals. [Lutes 2001](#) included all participants, with no loss of data for the outcomes of medication error, nosocomial infection, sepsis workups initiated, and thermal insults. However, for growth outcomes, study authors did not state the total number of participants analyzed nor reasons for loss of data.

Selective reporting

As illustrated in Table 1, [Chin 2006](#) and [Hayward 2007](#) provided evidence of selective outcome reporting.

Chin 2006 did not report the primary outcome of weight gain, as stated in the study title. Instead, study authors reported adjusted mean weight. From the study authors' primary data, we derived the rate of weight gain from data on weight, and we included this as the primary outcome in our analysis. Investigators did not assess length of hospital stay - a key outcome relevant to this study.

Hayward 2007 did not report the major pre-specified outcome of parental self efficacy and anxiety, nor the key outcomes of infant growth and apnea, bradycardia, and desaturation episodes. Study authors stated that "data were insufficient to analyse parental self-efficacy and anxiety."

Authors in Lutes 2001, Byers 2003a, Campbell-Yeo 2012b, and Badiie 2014 reported all pre-specified outcomes. However, Lutes 2001 did not provide the number of participants analyzed in any of their outcomes, nor standard deviations for continuous outcomes (mean weight gain, length, and head circumference growth).

Effects of interventions

See: [Summary of findings for the main comparison Co-bedding versus separate care for promoting growth and neurodevelopment in stable preterm twins](#)

Difficulties in combining data for meta-analysis

The six included studies presented their outcomes differently (Table 1). For instance, Campbell-Yeo 2012b reported A/B/D as a dichotomous outcome in the form of the number of infants in either group who had these episodes, but Chin 2006 reported this outcome as a continuous outcome in the form of the mean (and median) number of combined episodes in either group. Consequently, we could not combine data from these two studies. We included only data from Campbell-Yeo 2012b in RevMan (RevMan 2011) and analyzed the data of Chin 2006 separately (Stata 2007), as the data were skewed and were unsuitable for analysis in RevMan. Next, Byers 2003a reported only average figures for physiological measures such as heart rate, respiratory rate, and oxygen saturation - not the episodes of A/B/D specified a priori as our primary outcomes. Although Lutes 2001 reported outcomes reported in the other two studies, they provided combined data from twins and higher-order multiples and presented no separate report for twins only. Therefore, we could not extract appropriate data to be combined with data from the other studies. Furthermore, study authors did not report the standard deviation for continuous outcomes (weekly weight gain, gain in length, and head circumference). Among three studies that assessed caregiver errors, two documented no events (Byers 2003a; Hayward 2007), and Lutes 2001 documented one event in an infant in the co-bedding group. However, it remains unclear whether this infant was one of the twins or higher-order multiples - a population group that would have been excluded in this review. Therefore, we have not included the data for this outcome in our meta-analysis.

Overall, we were able to pool only two outcomes: PIPP in response to heel prick from Badiie 2014 and Campbell-Yeo 2012b, and infection from Byers 2003a, Chin 2006, and Hayward 2007. We combined all specific outcomes related to infection from each of the three studies above to form a composite outcome of

"Suspected or proven infection (any)." For all other outcomes, a single study effectively contributed the data: Chin 2006 provided data on weight; Hayward 2007 data on length of hospital stay, time in quiet sleep, crying, and co-regulation; Byers 2003a data on physiological measures such as heart rate, respiratory rate, oxygen saturation, and parental measures; and Campbell-Yeo 2012b infant pain measures and time to recovery of physiological parameters following painful procedures.

Adjustment of the unit of analysis

We obtained primary study data from the authors of Chin 2006 and adjusted the analysis by taking average measures of twin pairs for all outcomes, which included weight and total combined episodes of apnea, bradycardia, and desaturation (A/B/D). The authors of this paper did not report rate of weight gain, which we specified as the primary outcome of this review; instead, they reported "adjusted mean weight" as their primary outcome. We decided not to include "adjusted mean weight" as the outcome of our review, as we considered this misleading. Instead, we derived rate of weight gain (expressed as grams per kilogram of baseline weight per day) from the primary data provided by study authors and included this as the primary outcome of growth in this review. We used the formula $((\ln(\text{natural log})(\text{final weight}/\text{beginning weight}) \times 1000)/7 \text{ days})$ (Patel 2005) to derive the rate of weight gain for each week of the study from study entry to week three. We obtained the average rate of weight gain from study entry to week three by averaging the three readings that we obtained for each week of the study. As no information on length of hospital stay was included in the primary data, we could not include this in our analysis.

For Hayward 2007, we adjusted the unit of analysis from each individual infant to each twin pair for the outcomes of length of hospital stay, episodes of quiet sleep, crying, and co-regulation by halving the number of units (from six [individual infants] to three [twin pairs]). However, we could not adjust the analysis to overcome bias due to differential contributions of data from each twin pair for the outcomes of quiet sleep, crying, and co-regulation, as this would require the raw data that we have not managed to acquire. We also adjusted the unit of analysis for Byers 2003a from each individual infant to each twin pair, effectively halving the sample size. We performed no adjustment on the unit of analysis for Campbell-Yeo 2012b and Badiie 2014, as study authors had adjusted the results using GEE and reported no differences in effect sizes for their reported outcomes.

A summary of the effects of co-bedding versus separate care after adjustment for the unit of analysis follows here.

Primary outcomes

Rate of weight gain (in gram/kg of baseline weight/d) (Outcome 1.1)

We found no difference between co-bedded twins and twins receiving separate care in rate of weight gain throughout the period of the study (one study, 18 pairs of twins at the end of the study, low quality of evidence, downgraded two levels for study limitations and imprecision; see detailed results according to week after enrolment in Figure 4).

Figure 4. Forest plot of comparison: 1 Co-bedding vs separate care, outcome: 1.1 Rate of weight gain (in gram/kg of baseline weight/d).

Apnea, bradycardia, and desaturation (Outcome 1.2)

We found no difference in the number of infants with A/B/D in [Campbell-Yeo 2012b](#) (one study, 62 pairs of twins, RR 0.85, 95% CI 0.18 to 4.05, low quality of evidence, downgraded two levels for imprecision and indirectness). However, in [Chin 2006](#) (38 pairs of twins), the co-bedded group had consistently fewer episodes of A/B/D throughout the three weeks of the study compared with the control group, although the difference was not statistically significant ([Table 2](#)).

Co-regulated state (Outcome 1.3), crying (Outcome 1.4), and quiet sleep (Outcome 1.5)

Data from [Hayward 2007](#) showed no differences among episodes in the co-regulated state (MD 0.96, 95% CI -3.44 to 5.36, three pairs of twins, very low quality of evidence, downgraded three levels

because of imprecision and serious study limitations [two levels]). However, the co-bedded group had more crying episodes (one study, six pairs of twins, MD [scale of 0 to 20] 4.4, 95% CI 1.7 to 7.1) and more episodes of quiet sleep (one study, six pairs of twins, MD [scale of 0 to 20] 4.6, 95% CI 1.6 to 7.6).

Neurobehavior: infant pain score following painful procedure (Outcome 1.6)

Compared with twins who received separate care, co-bedded twins had lower mean PIPP score 30 seconds after heel prick (two studies, 117 pairs of twins, MD [scale of 0 to 21] -0.96, 95% CI -1.68 to -0.23, $I^2 = 75\%$, low quality of evidence, downgraded two levels for study limitations and inconsistency) but higher mean PIPP score 90 seconds after heel prick (one study, 62 pairs of twins, MD [scale of 0 to 21] 1.00, 95% CI 0.14 to 1.86) ([Figure 5](#)).

Figure 5. Forest plot of comparison: 1 Co-bedding vs separate care, outcome: 1.6 Neurobehavior: infant pain score following painful procedure.

For the outcome of PIPP score 30 seconds after heel prick, the two included studies differed in the direction of their results, with [Campbell-Yeo 2012b](#) showing no difference between the two groups and [Badiee 2014](#) showing marked differences favoring the co-bedded group. The I^2 statistic of 75% confirms substantial heterogeneity between study results.

Although we explored possible causes of the differences in results, we could not identify any plausible explanation in terms of risk of bias, population, nature of the intervention, and outcome measures. Both trials had low risk of bias in most domains, recruited infants of similar gestation and birth weight, and carried out co-bedding in a similar manner. Both studies evaluated pain using PIPP, a validated pain assessment scale for preterm infants.

The only major difference noted between the two studies was postnatal age of the infants when the study took place. In [Campbell-Yeo 2012b](#), investigators administered the interventions, which consisted of co-bedding of twins and heel prick, between 12 and 19 days of life, and [Badiee 2014](#) administered the intervention between days 3 and 4 of life. This difference in timing of the intervention might plausibly explain the heterogeneity on the basis of effects of co-bedding or infants' response to pain. First, the apparent lack of effectiveness of co-bedding observed in infants of greater postnatal age, as shown in [Campbell-Yeo 2012b](#), could be due to the fact that co-bedding was commenced after a longer period of separation since birth compared with [Badiee 2014](#). Alternatively, infants of greater postnatal age may have developed some form of adaptation to pain following prior exposure to painful procedures, and measures used at this time to alleviate pain, including co-bedding, might have become less effective than if used earlier. Notably, average PIPP scores in [Campbell-Yeo 2012b](#)

were lower than those in [Badiee 2014](#), supporting this postulation. Previous studies have consistently shown that newborn infants who had prior exposure to a greater number of painful procedures and those who stayed longer in the NICU had a dampened pain response ([Johnston 1999](#); [Grunau 2001](#); [Grunau 2005](#); [Williams 2009](#)). In [Campbell-Yeo 2012b](#), the average number of painful procedures in the co-bedded group and in the control group was 128.5 and 91.3, respectively, and although [Badiee 2014](#) did not report the number of painful procedures that infants had experienced before the start of the study, it was highly unlikely that a similar number of procedures would have been performed by day 3 or 4 of life. Nevertheless, the postulations above regarding the timing of co-bedding and infants' pain adaptation are hypotheses that await confirmation by future researchers.

After formulating some explanation for possible sources of heterogeneity, we decided to combine the results in a meta-analysis because the nature of the intervention, namely, co-bedding, was similar in both studies. If a greater number of studies are included for this outcome in the future, we will further evaluate our hypotheses by performing subgroup analyses based on the postnatal age of infants when the intervention was administered.

Secondary outcomes

Infection (Outcome 1.7)

We found no differences between the two groups in the overall incidence of suspected or proven infection (three studies, 65 pairs of twins, RR 0.84, 95% CI 0.30 to 2.31, very low quality of evidence, downgraded three levels for serious study limitations and imprecision) and in the incidence of specific infection, such as necrotizing enterocolitis, conjunctivitis, and sepsis ([Figure 6](#)).

Figure 6. Forest plot of comparison: 1 Co-bedding vs separate care, outcome: 1.7 Infections.

Length of hospital stay (Outcome 1.8)

Data from [Hayward 2007](#) reveal no differences in length of hospital stay between the two groups (one study, six pairs of twins, MD -4.9 days, 95% CI -35.23 to 25.43 days, very low quality of evidence, downgraded three levels for serious study limitations and imprecision).

Parental measures, including parental anxiety (Outcome 1.9), parental attachment (Outcome 1.10), and parental satisfaction (Outcome 1.11)

Data from [Byers 2003a](#) show no differences in all parental measures, including parental anxiety (one study, 18 pairs of twins, MD as measured by Parental State Anxiety Inventory [0 to 80, with higher score indicating higher anxiety level] 0.9, 95% CI -2.13 to 3.93), parental attachment (MD as measured by Maternal Attachment Inventory [0 to 112, with higher score indicating stronger attachment], one study, 18 pairs of twins, MD 0.9, 95% CI -2.02 to 3.82), and parental satisfaction (MD as measured by Parental Satisfaction Survey [0 to 55, with higher score indicating higher satisfaction level], one study, 18 pairs of twins, MD -0.38, 95% CI -4.49 to 3.73, moderate quality of evidence, downgraded one level for study limitations).

DISCUSSION

Summary of main results

In this review, a single study essentially contributed the data for most outcomes, as the included studies reported their outcomes differently. It was not possible to include the data from one study in the analysis. Very limited evidence showed no statistically significant differences between the co-bedded group and the group receiving separate care in terms of major outcomes, including rate of weight gain, A/B/D episodes, infection, episodes in co-regulated state, length of hospital stay, and parental perceptions of care. The co-bedded group appeared to have more crying episodes but seemed to spend more time in quiet sleep. The pain score for this group appeared to be lower 30 seconds after heel prick but slightly higher 90 seconds after heel prick compared with twins who received separate care, although marked heterogeneity in the results of pain score 30 seconds post heel prick precluded a clear conclusion. Overall major uncertainties on the benefits and harms of co-bedding preterm twins remain (see [Summary of findings for the main comparison](#) for details).

Overall completeness and applicability of evidence

We identified studies that included population, intervention, comparison, and outcomes that matched our pre-specified selection criteria, especially for our primary outcomes of growth (rate of weight gain), A/B/D episodes, and other major outcomes

of infection and parental perceptions of care. However, we found insufficient studies that examined neurodevelopment and other long-term outcomes. Lack of sufficient evidence on the benefits and harms of co-bedding, as our review has demonstrated, is reflective of the current state of practice, as co-bedding, although routinely practiced in some parts of the world, still is not widely practiced in neonatal nurseries worldwide.

Quality of the evidence

Overall low or very low quality of evidence for almost all outcomes assessed due to single-study contributions to each outcome and small sample sizes gave rise to imprecise estimates, heterogeneity, and limitations in the design of included studies, mainly as the result of non-blinding of outcome assessors, incomplete outcome data, and possible selective outcome reporting. Six studies addressed our research question but failed to meet the inclusion criteria in terms of study design. Overall, the body of evidence does not allow a clear conclusion to be made regarding the benefits and harms of co-bedding for stable preterm twins.

Potential biases in the review process

Strengths of this review include a comprehensive search of multiple databases, incorporation of additional study information not provided in published papers but obtained through contact with study authors, adjustments to the unit of analysis in data entry according to our pre-specified methods, and analysis based on raw data in one study. However, we could not obtain all the relevant data that we set out to obtain because of lack of response by the authors of one study; this could have introduced bias into our review.

Agreements and disagreements with other studies or reviews

One narrative review, which included non-randomized studies (Tomashek 2007), reported that available studies show that co-bedding did not improve physiological and/or clinical outcomes over separate care, or show improvements that were considered "not clinically significant." Study authors also commented that the overall quality of evidence was poor and recommended that additional research must be conducted. The findings of this narrative review are consistent with our findings.

AUTHORS' CONCLUSIONS

Implications for practice

Currently, evidence on the benefits and risks of co-bedding stable preterm twins is insufficient to permit recommendations for clinical practice.

Implications for research

The current low or very low quality of evidence was due to the small sample size of included studies, their limitations in design, and heterogeneity among study results. Randomized controlled trials with sufficient power are needed to detect clinically important differences by using twin pairs rather than individual infants as the unit of analysis, or by making statistical adjustments such as GEE to address unit of analysis issues, with clinically important outcomes such as rate of weight gain, neurobehavior, and neurodevelopment over the short term and over the long term. Future researchers should refer to existing evidence, such as that provided in systematic reviews that examine preterm infants' growth and neurodevelopment as primary outcomes, before they make a judgment on clinically important differences in those outcomes, to enable a reasonable estimate of the sample size required for future studies. If a study includes twins as well as higher-order multiples as participants, investigators should perform subgroup analyses of twins and higher-order multiples.

ACKNOWLEDGEMENTS

We thank Drs. Roger Soll, John Sinclair, Jeffrey Horbar, G Suresh, and Michael Bracken for their comments on the draft protocol and draft review. We thank Ms. Diane Haugthon, Managing Editor, and Ms. Yolanda Montage, Trials Search Co-ordinator of the Neonatal Review Group, for their assistance with tasks leading to publication of this review.

We are also grateful to the lead authors of our included studies - Susan DiNonno Chin and Kathryn Hayward - for generously providing additional information about their studies upon receipt of our requests.

REFERENCES

References to studies included in this review

Badiee 2014 {published data only}

Badiee Z, Nassiri Z, Armanian A. Cobedding of twin premature infants: calming effects on pain responses. *Pediatrics and Neonatology* 2014;**55**(4):262-8. [PUBMED: 24694748]

Byers 2003a {published data only}

Byers JF, Yovaish W, Lowman LB, Francis JD. Co-bedding versus single-bedding premature multiple-gestation infants in incubators. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 2003;**32**(3):340-7. [PUBMED: 12774876]

Campbell-Yeo 2012b {published data only}

Campbell-Yeo M, Johnston C, Chambers C, Joseph K, Feeley N, Barrington K. Do preterm twins mount an empathic response when being near their co-twin undergoing a painful procedure?. *The Journal of Pain* 2014;**15**(4):S31. [DOI: [10.1016/j.jpain.2014.01.127](https://doi.org/10.1016/j.jpain.2014.01.127)]

Campbell-Yeo ML, Johnston CC, Joseph K, Feeley NL, Chambers CT, Barrington KJ. Co-bedding as a comfort measure for twins undergoing painful procedures (CComForT Trial). *BMC Pediatrics* 2009;**9**:76. [PUBMED: 20003351]

Campbell-Yeo ML, Johnston CC, Joseph KS, Feeley N, Chambers CT, Barrington KJ. Co-bedding decreases time to recovery but does not diminish pain reactivity following heel lance in preterm twins (Abstract). *Pain Research and Management* 2012; Vol. 17, issue 3:207.

* Campbell-Yeo ML, Johnston CC, Joseph KS, Feeley N, Chambers CT, Barrington KJ. Cobedding and recovery time after heel lance in preterm twins: results of a randomized trial. *Pediatrics* 2012;**130**(3):500-6. [PUBMED: 22926182]

Campbell-Yeo ML, Johnston CC, Joseph KS, Feeley N, Chambers CT, Barrington KJ, et al. Co-bedding between preterm twins attenuates stress response after heel lance: results of a randomized trial. *Clinical Journal of Pain* 2014;**30**(7):598-604. [PUBMED: 24300226]

Chin 2006 {published data only}

Chin SD, Hope L, Christos PJ. Randomized controlled trial evaluating the effects of cobedding on weight gain and physiologic regulation in preterm twins in the NICU. *Advances in Neonatal Care* 2006;**6**(3):142-9. [PUBMED: 16750808]

Hayward 2007 {published data only}

Hayward K, Campbell-Yeo M, Price S, Morrison D, Whyte R, Cake H, et al. Co-bedding twins: how pilot study findings guided improvements in planning a larger multicenter trial. *Nursing Research* 2007;**56**(2):137-43. [PUBMED: 17356445]

Lutes 2001 {published data only}

Lutes LM, Altimier L. Co-bedding multiples. *Newborn & Infant Nursing Reviews* 2001;**1**(4):242-6.

References to studies excluded from this review

Anon 2002 {published data only}

Anon. Cobedding low birth weight twins. *Nursing Australia* 2002;**3**(1):5.

DellaPorta 1998 {published data only}

DellaPorta K, Aforismo D, Butler-O'Hara M. Co-bedding of twins in the neonatal intensive care unit. *Pediatric Nursing* 1998;**24**(6):529-31. [PUBMED: 10085994]

Fischer 1991 {published data only}

Fischer PR, Dind Y. Co-sleeping and neonatal weight loss. *Annals of Tropical Paediatrics* 1991;**11**(2):189-91. [PUBMED: 1715152]

Jahanfar 2012 {published data only}

Jahanfar S. Twins co-bedding at home, parents' perspective, sleeping pattern and developmental milestones. *Shiraz E-Medical Journal* 2012;**13**(1):13-8.

LaMar 2006 {published data only}

LaMar K, Dowling DA. Incidence of infection for preterm twins cared for in cobedding in the neonatal intensive-care unit. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 2006;**35**(2):193-8. [PUBMED: 16620244]

Longobucco 2002 {published data only}

Longobucco D, Bernstein B, Rossi D. To co-bed or not to co-bed? A comparative study of co-bedded versus individually bedded multiple birth infants in the NICU. *Pediatric Research* 2002;**47**:413S.

Lutes 2000 {published data only}

Lutes LM, Altimier L. Co-bedding twins and higher-order multiples. *Central Lines* 2000;**16**(2):10-3.

Matthews 2004 {published data only}

Matthews T, McDonnell M, McGarvey C, Loftus G, O'Regan M. A multivariate "time based" analysis of SIDS risk factors. *Archives of Disease in Childhood* 2004;**89**(3):267-71. [PUBMED: 14977707]

Mazeiras 2010 {published data only}

Mazeiras G, des Robert C, Legrand A, Caillaux G, Corre A, Roze JC, et al. Cobedding for premature twins [Le cobedding pour les jumeaux prématures]. *Soins Pédiatrie, Puericulture* 2010;**31**(254):20-2.

Nyqvist 1998 {published data only}

Nyqvist KH, Lutes LM. Co-bedding twins: a developmentally supportive care strategy. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 1998;**27**(4):450-6. [PUBMED: 9684208]

Nyqvist 2002 {published data only}

Nyqvist KH. Breast-feeding in preterm twins: development of feeding behavior and milk intake during hospital stay and related care giving practices. *Journal of Pediatric Nursing* 2002;**17**(4):246-56. [PUBMED: 12219324]

Polizzi 2003 {published data only}

Polizzi J, Byers JF, Kiehl E. Co-bedding versus traditional bedding of multiple-gestation infants in the NICU. *Journal for Healthcare Quality* 2003;**25**(1):5-10; quiz 10-1. [PUBMED: 12879624]

Stainton 2005 {published data only}

Stainton C, Jozsa E, Fethney J. Responses to co-bedding of low birth weight twins in neonatal intensive care. *Neonatal, Paediatric and Child Health Nursing* 2005;**8**:4-12.

Touch 2002 {published data only}

Touch SM, Epstein ML, Pohl CA, Greenspan JS. The impact of cobedding on sleep patterns in preterm twins. *Clinics in Pediatrics* 2002;**41**(6):425-31. [PUBMED: 12166795]

Wahl 2006 {published data only}

Wahl SA. Co-sleeping and baby-controlled breast feeding [Samsoving og barnestyrte amming]. *Tidsskrift for den Norske Lægeforening* 2006;**126**(17):2282. [PUBMED: 16967070]

References to ongoing studies
NCT01480856 {published and unpublished data}

Co-bedding in Daily Weight Gain of Neonate Twins. Ongoing study September 2008.

Additional references
Adegbite 2004

Adegbite AL, Castille S, Ward S, Bajoria R. Neuromorbidity in preterm twins in relation to chorionicity and discordant birth weight. *American Journal of Obstetrics and Gynecology* 2004;**190**(1):156-63. [PUBMED: 14749653]

Adler 2006

Adler MR, Hyderi A, Hamilton A, Brown P. Clinical inquiries: what are safe sleeping arrangements for infants?. *Journal of Family Practice* 2006;**55**(12):1083-4, 1087. [PUBMED: 17137548]

American Academy of Pediatrics 2003

American Academy of Pediatrics. Committee of Fetus and Newborn. Apnea, sudden infant death syndrome, and home monitoring. *Pediatrics* 2003;**111**(4 Pt 1):914-7. [PUBMED: 12671135]

American Academy of Pediatrics 2005

American Academy of Pediatrics Task Force on Sudden Infant Death Syndrome. The changing concept of sudden infant death syndrome: diagnostic coding shifts, controversies regarding the sleeping environment, and new variables to consider in reducing risk. *Pediatrics* 2005;**116**(5):1245-55. [PUBMED: 16216901]

Bagchi 2006

Bagchi S, Salihi HM. Birth weight discordance in multiple gestations: occurrence and outcomes. *Journal of Obstetrics and Gynaecology* 2006;**26**(4):291-6. [PUBMED: 16753674]

Blondel 2002

Blondel B, Kaminski M. Trends in the occurrence, determinants, and consequences of multiple births. *Seminars in Perinatology* 2002;**26**(4):239-49. [PUBMED: 12211614]

Blondel 2002a

Blondel B, Kogan MD, Alexander GR, Dattani N, Kramer MS, Macfarlane A, et al. The impact of the increasing number of multiple births on the rates of preterm birth and low birthweight: an international study. *American Journal of Public Health* 2002;**92**(8):1323-30. [PUBMED: 12144992]

Bonellie 2005

Bonellie SR, Currie D, Chalmers J. Comparison of risk factors for cerebral palsy in twins and singletons. *Developmental Medicine and Child Neurology* 2005;**47**(9):587-91. [PUBMED: 16138664]

Branum 2003

Branum AM, Schoendorf KC. The effect of birth weight discordance on twin neonatal mortality. *Obstetrics and Gynecology* 2003;**101**(3):570-4. [PUBMED: 12636964]

Bredemeyer 2012

Bredemeyer SL, Foster JP. Body positioning for spontaneously breathing preterm infants with apnoea. *Cochrane Database of Systematic Reviews* 2012, Issue 6. [DOI: [10.1002/14651858.CD004951.pub2](https://doi.org/10.1002/14651858.CD004951.pub2)]

Buscher 2000

Buscher U, Horstkamp B, Wessel J, Chen FC, Dudenhausen JW. Frequency and significance of preterm delivery in twin pregnancies. *International Journal of Gynaecology and Obstetrics* 2000;**69**(1):1-7. [PUBMED: 10760526]

Byers 2003b

Byers JF. Components of developmental care and the evidence for their use in the NICU. *American Journal of Maternal Child Nursing* 2003;**28**(3):174-80; quiz 181-2. [PUBMED: 12771696]

Cheung 1995

Cheung VY, Bocking AD, Dasilva OP. Preterm discordant twins: what birth weight difference is significant?. *American Journal of Obstetrics and Gynecology* 1995;**172**(3):955-9. [PUBMED: 7892890]

Eriksson 2007

Eriksson AW, Fellman J. Temporal trends in the rates of multiple maternities in England and Wales. *Twin Research and Human Genetics* 2007;**10**(4):626-32. [PUBMED: 17708703]

Feldman 2002

Feldman R, Eidelman AI, Sirota L, Weller A. Comparison of skin-to-skin (kangaroo) and traditional care: parenting outcomes and preterm infant development. *Pediatrics* 2002;**110**(1 Pt 1):16-26. [PUBMED: 12093942]

Gilbert 2005

Gilbert R, Salanti G, Harden M, See S. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of

recommendations from 1940 to 2002. *International Journal of Epidemiology* 2005;**34**(4):874-87. [PUBMED: 15843394]

Glinianaia 2006

Glinianaia SV, Jarvis S, Topp M, Guillem P, Platt MJ, Pearce MS, et al. Intrauterine growth and cerebral palsy in twins: a European multicenter study. *Twin Research and Human Genetics* 2006;**9**(3):460-6. [PUBMED: 16790158]

Goyen 2003

Goyen TA, Veddovi M, Lui K. Developmental outcome of discordant premature twins at 3 years. *Early Human Development* 2003;**73**(1-2):27-37. [PUBMED: 12932891]

GRADEpro [Computer program]

McMaster University. GRADEpro [www.gradeapro.org]. McMaster University, 2014.

Gray 2003

Gray PH, Flenady V. Cot-nursing versus incubator care for preterm infants. *Cochrane Database of Systematic Reviews* 2003, Issue 1. [DOI: [10.1002/14651858.CD003062.pub2](https://doi.org/10.1002/14651858.CD003062.pub2)]

Grunau 2001

Grunau RE, Oberlander TF, Whitfield MF, Fitzgerald C, Lee SK. Demographic and therapeutic determinants of pain reactivity in very low birth weight neonates at 32 weeks' postconceptional age. *Pediatrics* 2001;**107**:105-12.

Grunau 2005

Grunau RE, Holsti L, Haley DW, Oberlander T, Weinberg J, Solimano A, et al. Neonatal procedural pain exposure predicts lower cortisol and behavioral reactivity in preterm infants in the NICU. *Pain* 2005;**113**(3):293-300. [PUBMED: 15661436]

Gutbrod 2000

Gutbrod T, Wolke D, Soehne B, Ohrt B, Riegel K. Effects of gestation and birth weight on the growth and development of very low birthweight small for gestational age infants: a matched group comparison. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2000;**82**(3):F208-14. [PUBMED: 10794788]

Guyatt 2011a

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2011;**64**(4):383-94. [PUBMED: 21195583]

Guyatt 2011b

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence - study limitations (risk of bias). *Journal of Clinical Epidemiology* 2011;**64**(4):407-15. [PUBMED: 21247734]

Guyatt 2011c

Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence - imprecision. *Journal of Clinical Epidemiology* 2011;**64**(12):1283-93. [PUBMED: 21839614]

Guyatt 2011d

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence - inconsistency. *Journal of Clinical Epidemiology* 2011;**64**(12):1294-302. [PUBMED: 21803546]

Guyatt 2011e

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence - indirectness. *Journal of Clinical Epidemiology* 2011;**64**(12):1303-10. [PUBMED: 21802903]

Hack 2003

Hack M, Schluchter M, Cartar L, Rahman M, Cuttler L, Borawski E. Growth of very low birth weight infants to age 20 years. *Pediatrics* 2003;**112**(1 Pt 1):e30-8. [PUBMED: 12837903]

Hayward 2003

Hayward K. Cobedding of twins: a natural extension of the socialization process?. *American Journal of Maternal Child Nursing* 2003;**28**(4):260-3. [PUBMED: 12840693]

Henderson-Smart 1986

Henderson-Smart DJ, Butcher-Puech MC, Edwards DA. Incidence and mechanism of bradycardia during apnoea in preterm infants. *Archives of Disease in Childhood* 1986;**61**(3):227-32. [PUBMED: 3963865]

Higgins 2011

Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1. 0 [updated March 2011]. The Cochrane Collaboration, 2009. www.cochrane-handbook.org.

Hur 2005

Hur YM, Kwon JS. Changes in twinning rates in South Korea: 1981-2002. *Twin Research and Human Genetics* 2005;**8**(1):76-9. [PUBMED: 15836815]

Imaizumi 1998

Imaizumi Y. A comparative study of twinning and triplet rates in 17 countries, 1972-1996. *Acta Geneticae Medicae et Gemellologiae (Roma)* 1998;**47**(2):101-14. [PUBMED: 10783768]

Jacquemyn 2003

Jacquemyn Y, Martens G, Ruysinck G, Michiels I, Van Overmeire B. A matched cohort comparison of the outcome of twin versus singleton pregnancies in Flanders, Belgium. *Twin Research* 2003;**6**(1):7-11. [PUBMED: 12626222]

Johnston 1999

Johnston CC, Stevens BJ, Franck LS, Jack A, Stremmer R, Platt R. Factors explaining lack of response to heel stick in preterm newborns. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 1999;**28**(6):587-94. [PUBMED: 10584912]

Lester 2004

Lester BM, Tronick EZ, Brazelton TB. The Neonatal Intensive Care Unit Network Neurobehavioral Scale procedures. *Pediatrics* 2004;**113**(3 Pt 2):641-67. [PUBMED: 14993524]

Liu 2002

Liu YC, Blair EM. Predicted birthweight for singletons and twins. *Twin Research* 2002;**5**(6):529-37. [PUBMED: 12573184]

Moore 2012

Moore ER, Anderson GC, Bergman N, Dowswell T. Early skin-to-skin contact for mothers and their healthy newborn infants. Cochrane Database of Systematic Reviews. John Wiley & Sons, Ltd, 2012, issue 5. [DOI: [10.1002/14651858.CD003519.pub3](https://doi.org/10.1002/14651858.CD003519.pub3)]

Patel 2005

Patel AL, Engstrom JL, Meier PP, Kimura RE. Accuracy of methods for calculating postnatal growth velocity for extremely low birth weight infants. *Pediatrics* 2005;**116**(6):1466-73. [PUBMED: 16322172]

Pharoah 2007

Pharoah PO, Platt MJ. Sudden infant death syndrome in twins and singletons. *Twin Research and Human Genetics* 2007;**10**(4):644-8. [PUBMED: 17708706]

RevMan 2011 [Computer program]

The Nordic Cochrane Centre. The Cochrane Collaboration. Review Manager (RevMan). Version 5.1. Copenhagen: The Nordic Cochrane Centre. The Cochrane Collaboration, 2011.

Scher 2002

Scher AI, Petterson B, Blair E, Ellenberg JH, Grether JK, Haan E, et al. The risk of mortality or cerebral palsy in twins: a collaborative population-based study. *Pediatric Research* 2002;**52**(5):671-81. [PUBMED: 12409512]

Schünemann 2013

Schünemann H, Brożek J, Guyatt G, Oxman A, editors. GRADE WG. GRADE handbook for grading quality of evidence and strength of recommendations. www.guidelinedevelopment.org/handbook. Updated October 2013.

Stata 2007 [Computer program]

StataCorp. Stata Statistical Software. Version Release 10. College Station, TX: StataCorp, 2007.

Symington 2006

Symington A, Pinelli J. Developmental care for promoting development and preventing morbidity in preterm infants. *Cochrane Database of Systematic Reviews* 2006, Issue Issue 2. [DOI: [10.1002/14651858.CD001814.pub2](https://doi.org/10.1002/14651858.CD001814.pub2)]

Tomashek 2007

Tomashek KM, Wallman C, Committee on Fetus and Newborn, American Academy of Pediatrics. Cofbedding twins and

higher-order multiples in a hospital setting. *Pediatrics* 2007;**120**(6):1359-66. [PUBMED: 18055686]

Topp 2004

Topp M, Huusom LD, Langhoff-Roos J, Delhumeau C, Hutton JL, Dolk H. Multiple birth and cerebral palsy in Europe: a multicenter study. *Acta Obstetrica et Gynecologica Scandinavica* 2004;**83**(6):548-53. [PUBMED: 15144336]

Washington 1998

Washington K. The Bayley Scales of Infant Development-II and children with developmental delays: a clinical perspective. *Journal of Developmental and Behavioral Pediatrics* 1998;**19**(5):346-9. [PUBMED: 9809265]

WHO 1997

Department of Reproductive Health and Research (RHR), World Health Organization. Thermal Protection of the Newborn: A Practical Guide. Geneva: World Health Organization, 1997.

Williams 2009

Williams AL, Khattak AZ, Garza CN, Lasky RE. The behavioral pain response to heelstick in preterm neonates studied longitudinally: description, development, determinants, and components. *Early Human Development* 2009;**85**(6):369-74. [PUBMED: 19167172]

Yalcin 1998

Yalcin HR, Zorlu CG, Lembet A, Ozden S, Gokmen O. The significance of birth weight difference in discordant twins: a level to standardize?. *Acta Obstetrica et Gynecologica Scandinavica* 1998;**77**(1):28-31. [PUBMED: 9492713]

Yinon 2005

Yinon Y, Mazkereth R, Rosentzweig N, Jarus-Hakak A, Schiff E, Simchen MJ. Growth restriction as a determinant of outcome in preterm discordant twins. *Obstetrics and Gynecology* 2005;**105**(1):80-4. [PUBMED: 15625146]

References to other published versions of this review

Lai 2012

Lai NM, Foong S Cheng, Foong W Cheng, Tan K. Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins. *Cochrane Database of Systematic Reviews* 2012, Issue 12. [DOI: [10.1002/14651858.CD008313.pub2](https://doi.org/10.1002/14651858.CD008313.pub2)]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Badiee 2014

Methods	Single-centre RCT in a NICU at a university hospital (Iran)
---------	---

Badiee 2014 (Continued)

Participants	105 newborn infants with gestational age of 26 to 34 weeks and postnatal age of less than 20 days who underwent heel blood sampling for blood glucose determination. Study authors did not state the reasons for the odd number recruited. 100 of these infants (50 pairs of twins) received the allocated intervention. Infants who had received sedatives or analgesics, with major congenital malformations, an Apgar scores < 6 at 5 minutes after birth, severe respiratory distress requiring mechanical ventilation, or severe intraventricular haemorrhage were excluded
Interventions	Co-bedding (cared for in the same incubator) vs separate care (cared for in separate incubators) with an aim to assess their pain response to heel prick. "Newborns in the co-bedding group were placed side by side in an incubator without any clothing except for diapers so that they could touch each other freely, with each side of the incubator pertaining to one twin. Infants were co-bedded from 24 hours prior to heel sticks to the end of the study" Although not stated clearly, it appeared that heel prick was performed for all infants
Outcomes	Infant pain response to heel prick, assessed by the Premature Infant Pain Profile (PIPP). Secondary outcome was the amount of salivary cortisol secreted after the heel prick procedure
Notes	Other parameters assessed were duration of crying and highest heart rate during heel prick, which together with the amount of salivary cortisol were not included in our meta-analysis, as they were not included as pre-specified outcomes of our review

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Materials and methods: "Randomisation was performed using a computer-generated random number algorithm"
Allocation concealment (selection bias)	Low risk	Materials and methods: "Allocation of eligible newborns to the intervention and the control groups was performed using a sealed opaque envelope"
Incomplete outcome data (attrition bias) Sepsis	Low risk	100 out of 105 infants recruited (95.2%) received the allocated intervention, all of whom were included in the subsequent analysis Five infants (number of twin sets unclear) did not receive co-bedding as originally allocated and were excluded from the analysis with no reasons stated
Selective reporting (reporting bias)	Low risk	Both major outcomes stated in the methods, namely, PIPP and salivary cortisol levels, were reported in sufficient detail in the results
Other bias	Low risk	None identified
Blinding of participants and personnel (performance bias) All outcomes	High risk	Study authors stated, "Blood sampling was performed in a standardized manner by expert technicians who could not be blinded to the study" (Materials and methods). Blinding of the other caregivers was highly unlikely in view of differences between the 2 allocated groups in the way the infants were cared for
Blinding of outcome assessment (detection bias) All outcomes	High risk	Study authors stated that "the researchers could not be blinded for the assigned groups"

Byers 2003a

Methods	Randomized repeated measures study at a tertiary level NICU (USA)
---------	---

Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins (Review)

Byers 2003a (Continued)

Participants	A convenience sample of 37 preterm infants (average 29 weeks' PMA) and 19 parents. Exclusion criteria "included the infants having a known infection, being on mechanical ventilation, and having a combined weight too great to be safely co-bedded in an incubator." Study authors did not specify the combined weight limit
Interventions	Co-bedding (cared for in the same incubator) vs separate care (cared for in separate incubators). Study authors stated that only 2 infants were co-bedded at a time. For triplets, the first 2 infants to achieve physiological stability as determined by the clinicians involved were co-bedded. In total, 16 infants (8 pairs) were co-bedded and 21 infants were assigned to separate care. It was unclear how many of these infants were twins and how many were a part of triplets
Outcomes	Main outcomes were physiological parameters, sleep-wake synchrony of infants, neurobehavioral observations using the Newborn Individualized Developmental Care and Assessment Program (NIDCAP) guidelines, and parental measures including anxiety (measured using Parental State Anxiety Inventory, an 80-point scale, with higher score indicating higher anxiety level), attachment (measured using Maternal Attachment Inventory, a 112-point scale, with higher score indicating higher attachment), and satisfaction (measured using Parental Satisfaction Survey, a 55-point scale, with higher score indicating higher satisfaction level)
Notes	<ul style="list-style-type: none"> Study authors stated, "33 of the participating infants were twins, and four were triplets." They did not provide an explanation of the numbers of twins and triplets quoted above that did not total up to the appropriate numbers for twins and triplets, respectively. In particular, study authors did not explain how many sets of twins and triplets were enrolled and how many "incomplete sets" actually participated Study authors stated, "Single-bedded infants received the traditional NICU standard of care. Co-bedded infants received care using the institution's co-bedding care protocol," which involved details of infant care, monitoring, and hand hygiene. It was unclear whether infants in the control group were cared for according to the same protocol as adapted for individual infants This article reported average figures for physiological measures such as heart rate, respiratory rate, and oxygen saturation - not episodes of apnea, bradycardia, or desaturation, which we have specified a priori as our primary outcomes Unit of analysis issue was that all outcome data were analyzed and reported for each individual infant rather than for each twin pair The Maternal Attachment Inventory and the State-Traits Anxiety Inventory, used to measure parental attachment and anxiety, respectively, had poor reliability overall, with Cronbach's alpha measuring as low as 0.44 and 0.54, respectively It was unclear how one of the outcomes - "daily breast and/or formula milk ingestion (ml)" - was measured, specifically, how the amount of breast milk ingested was recorded. It was not stated whether all breast-feeding infants ingested expressed breast milk during the study

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Participants and sample: "The participants were assigned randomly." No other statements on how random sequence was generated were available
Allocation concealment (selection bias)	Unclear risk	Participants and sample: "The participants were assigned randomly." No further statements indicated whether random sequence generation was performed independently from allocation, or how allocation was carried out
Incomplete outcome data (attrition bias) Sepsis	Low risk	All except 1 infant in the control group completed the study. This single infant was withdrawn from the study after sepsis was diagnosed

Byers 2003a (Continued)

Selective reporting (reporting bias)	Low risk	All outcomes pre-specified in the methods were reported in sufficient detail in the results
Other bias	High risk	Some imbalance in major characteristics was noted between the 2 assigned groups. Co-bedded infants had higher PMA at study enrollment than infants in the control group (average 33 weeks PMA in co-bedding group, 30.6 weeks PMA in control group), and a greater proportion of infants were male (68.3% in the co-bedded group, 22.7% in the control group)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Although not stated in the paper, blinding for caregivers and for parents appeared highly unlikely. Study authors stated, "Single-bedded infants received the traditional NICU standard of care. Co-bedded infants received care using the institution's co-bedding care protocol," which involved details of infant care, monitoring, and hand hygiene. It was unclear whether the "traditional NICU standard of care" as stated followed the same protocol as the co-bedded group, other than necessary differences that arose from caring for 1 infant vs 2 in an incubator or crib
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	It was not stated whether the data collector was blinded to the status of the infants

Campbell-Yeo 2012b

Methods	Multi-centre randomized controlled trial involving 3 tertiary-level university-affiliated NICUs in Canada	
Participants	Study setting and population: "Twins were considered eligible for recruitment if they were heavier than 1000 g, without major anomalies, requiring at least 1 medically indicated heel lance for blood procurement, and considered medically stable (without infection, indwelling chest tubes or umbilical catheters, or need for mechanical ventilation). Twins were not considered eligible if only 1 twin required overhead phototherapy at time of the heel lance." A total of 67 twin sets (134 infants) were randomized: 36 sets (72 infants) to the co-bedding group and 31 sets (62 infants) to the control group	
Interventions	Co-bedding (cared for in the same incubator or crib) vs "separate care" (cared for in separate incubators or cribs). All participants also received 1 dose of 24% sucrose 2 minutes before the heel lance and were offered a pacifier. Additional (rescue) doses of 24% sucrose were offered as deemed necessary by the nurse caring for the twins. as per usual care practices	
Outcomes	The primary outcome was pain response as measured by the Premature Infant Pain Profile (PIPP), a previously validated 21-point scale, with higher scores indicating greater pain. Other outcomes included time required for physiological recovery in response to heel lance, determined by length of time for heart rate and oxygen saturation to return to baseline, the need for additional pain relief with 24% sucrose, and adverse outcomes, such as episodes of apnea, bradycardia, infection. and caregiver error	
Notes	Study authors adjusted for non-dependence between twin pairs by using the generalized estimating equation, and for differences in baseline characteristics by using regression techniques (statistical analysis)	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Study design and intervention:

Campbell-Yeo 2012b (Continued)

		Participants were randomized after parental consent by "randomly permuted blocks of 2, 4, or 6, via a computerised off-site Web site accessed by the principal investigator (M.C-Y.) or research nurse"
Allocation concealment (selection bias)	Low risk	Study design and intervention: Participants were randomized after parental consent by "randomly permuted blocks of 2, 4, or 6, via a computerised off-site Web site accessed by the principal investigator (M.C-Y.) or research nurse"
Incomplete outcome data (attrition bias) Sepsis	Low risk	A total of 10 infants (5 from each group) out of 134 did not complete the study. The number of non-completers was small and was balanced between the 2 groups; reasons for non-completion (transfer before heel prick and physiological data equipment malfunction) were unlikely to be related to outcomes
Selective reporting (reporting bias)	Low risk	All outcomes pre-specified in the methods were reported in sufficient detail in the results
Other bias	Low risk	None identified
Blinding of participants and personnel (performance bias) All outcomes	High risk	It was not stated whether caregivers and parents were blinded to group assignment, although this appeared highly unlikely
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	It was not stated whether the research nurse who collected data on pain score and physiological parameters was blinded to the status of the infants. One statement suggested that the data analyst might be blinded (Statistical analysis: "All data were analysed in a research laboratory off-site from the NICU")

Chin 2006

Methods	Randomized controlled trial at a tertiary-level NICU (USA)
Participants	Thirty-nine sets of preterm twins born at less than 34 weeks' PMA, who were clinically stable at the time of enrollment
Interventions	Co-bedding (cared for in the same incubators) vs separate care (cared for in separate incubators). A protocol was drawn for co-bedding twins enrolled in the study. In the co-bedding group, twins were placed side by side, initially in an incubator and later in an open crib, when both reached 1700 grams in weight. The same nurse looked after both twins. One blanket was used to lightly swaddle the twins together, allowing the hands to move freely to enable physical contact between the twins. Twins were separated if one or both became unstable. For twin sets in the control group, the only statements concerning their care were these: "The twin sets that were randomised to the control group received routine care in separate beds" (Purpose and methods, paragraph 2), and "Twin sets in both groups were assigned to the same bedside nurse" (Purpose and methods, paragraph 4)
Outcomes	Weight and apnea, bradycardia, and desaturation (A/B/D) episodes were collected from electronic medical records. Apnea was defined by study authors as cessation of breathing for > 20 seconds, bradycardia as any heart rate < 80 beats per minute, and desaturation as any oxygen saturation level < 85% (Purpose and methods, paragraph 6). If any of these events occurred singularly, they were counted as an individual event. If more than 1 type of event occurred in a cluster, they were counted as a single event Study authors reported adjusted mean weights (after controlling for baseline weights) at weeks 1, 2, and 3, and median number of combined A/B/D episodes

Chin 2006 (Continued)

Notes

- Unit of analysis issue: All outcome data are analyzed and reported as for each individual infant instead of for each twin pair
- Data for outcomes of combined A/B/D episodes were skewed and were not analyzed in RevMan. We used a primary statistical analysis software ([Stata 2007](#)) to analyze the data

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The method of random sequence generation was not mentioned. Correspondence with study authors revealed that the randomisation list was manually created by the first 2 authors, but no detail was provided on the tools used
Allocation concealment (selection bias)	Low risk	"Twins were randomised using sealed envelopes that stated either control or experimental group" (Purpose and methods). Correspondence with study authors revealed that sequence allocation was carried out by using "opaque and sealed envelopes"
Incomplete outcome data (attrition bias) Sepsis	Low risk	<p>Study authors mentioned reasons for withdrawal from the study. However, the sample became progressively smaller over the 3 weeks of the study, with loss of participants greater than could be accounted for by withdrawal. Attrition rates differed between the 2 groups over the duration of the study, as by the second week, 13 pairs of twins in the co-bedding group and 19 pairs in the control groups remained, and by the third week, 10 pairs of twins in the co-bedding group and 12 pairs in the control group remained. Reasons for this were not directly given but could be inferred from the following statement in Results, paragraph 4: "The data were analysed for the first 3 weeks of enrolment. After this time, the sample size became too small due to discharge home from the hospital"</p> <p>Study authors also explained the reasons why data collection was withheld for some participants during the course of the study, as follows (Results, paragraph 3): "In the co-bedded group, twins were separated due to complications consisting of conjunctivitis (1 set separated for 2 days); PDA ligation (1 set separated for 1 week); co-bedding and data collection were resumed once the infants were considered stable based on study guidelines"</p>
Selective reporting (reporting bias)	High risk	<ul style="list-style-type: none"> • Although the primary outcome of the study was assessed weight gain (Title and Purpose and methods, paragraph 1), weight gain was not directly reported. Instead, study authors reported baseline weights and weights at each point of measurement (weeks 1, 2, and 3), adjusted for differences in baseline weight • Length of hospital stay was not reported. This was an important outcome in view of differences in the number of twin sets discharged home over the 3 weeks of the study, as highlighted above under "Incomplete outcome data assessment" <p>Table 1 provides a matrix on the outcomes listed under methods and any additional key outcomes and actual outcomes reported</p>
Other bias	Low risk	None identified
Blinding of participants and personnel (performance bias) All outcomes	High risk	Although not stated in the article, blinding of caregivers and parents appeared highly unlikely

Chin 2006 (Continued)

Blinding of outcome assessment (detection bias) All outcomes	High risk	Although not stated in the paper, further correspondence with the lead study author confirmed that the research assistant who collected the data was not blinded to the status of the twin pairs
---	-----------	--

Hayward 2007

Methods	Pilot randomized controlled trial at a tertiary-level NICU (Canada)	
Participants	Six pairs of twins < 2000 grams. It was not stated whether weight criteria referred to weight at birth or at study entry	
Interventions	Co-bedding (cared for in the same incubator) vs separate care (cared for in separate incubators). Twin sets in the co-bedding group were placed in a common twin incubator (Giraffe, GE Medical, Fairfield, CT) for care. Twins in the control group received standard neonatal care in separate incubators. Each twin set was cared for by the same nurse during each shift. Data for each twin set were recorded over 8 blocks of observational periods over 2 weeks, each lasting 5 hours	
Outcomes	<p>Parental self efficacy, parental anxiety, infant "co-regulation" (defined as synchrony of infant state, heart rate, respirations, temperature, and oxygen saturation - regardless of whether the twins were co-bedded), infection rate, and incidence of caregiver error</p> <p>A research assistant recorded infants' physiological parameters from their charts and transcribed data onto the Nursing Child Assessment Sleep/Activity Record (NCASAR) for comparison of the record between twin dyads for co-regulatory activities. In the article, study authors did not define what thresholds were used to define the presence of "co-regulation" nor was the unit of measurement used for the outcomes of "time in quiet sleep," "time crying," and "co-regulation" stated. Infection rate was determined via infants' charts and medical records, by 3 measures: incidence of septic workup, treatment with antibiotics, and confirmed incidence of sepsis. Caregiver errors were collected from the nursing quarterly report, rather than recorded through direct observation. Infants were observed for 2 weeks</p> <p>Study authors provided more detailed information through further correspondence on the unit of measurement for the outcomes of "time in quiet sleep," "time crying," and "co-regulation," as follows:</p> <p>"The same unit of measurement was used for all outcomes. Each 15 minutes for a five hour block of time, infants were observed for 5 minutes. The research assistant would then assign their state for that 5 minute block of time. The outcome refers to the number of times infants were assigned a specific state. Coregulation was assigned if the subjects (twin set) were in the same state or were transitioning to the same state. Transitioning was defined as being only one state apart (e.g., one twin quiet sleep and one twin active sleep, or one twin drowsy and one twin quiet alert)"</p> <p>This essentially translated into a measurement scale of 0 to 20, and the rating reflected the number of times a twin pair was observed to have the outcome of interest during the 20 observation periods scheduled within the 5-hour study period. A minimum of 8 five-hour blocks were completed for each twin pair on a biweekly basis. All ratings in each assigned group were averaged to obtain final estimates</p>	
Notes	Unit of analysis issue was that all outcome data were analyzed and reported for each individual infant rather than for each twin pair	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Twins were randomised into two groups" (Design, paragraph 1). No further mention was made of randomization methods in the paper. Study author provided the following additional information: "Twins were randomised using a simple card draw. Randomisation assignment was written on a card and placed in an envelope. The [principal] investigator placed equal assignment

Hayward 2007 (Continued)

		cards (7 co-bedding and 7 separate care) into 14 envelopes. These envelopes were placed in a box"
Allocation concealment (selection bias)	Low risk	"Group allocation was assigned randomly using opaque envelopes" (Sample, paragraph 1). Study author provided the following additional information: "The randomisation assignment was placed on the inside of a folded card. This card was then taped shut and placed in a sealed envelope" and "The research assistant blindly selected an envelope from a box after the consent was obtained"
Incomplete outcome data (attrition bias) Sepsis	High risk	<p>It was unclear how many pairs of twins were analyzed for each outcome, and loss of data was not mentioned. Further correspondence with the lead study author revealed that complete data for the whole 8 periods of observation were available for only 2 sets of twins. Another 2 sets of twins were observed for 7 out of 8 periods; 1 set was observed for 4 periods, and another for 2 periods. Reasons for inability to obtain complete outcome data were stated: "Due to early discharge or unplanned transfer to home hospital, not all twin sets had eight five-hour blocks of analysis"</p> <p>In their final estimates for the outcomes of "time in quiet sleep," "time crying," and "co-regulated," study authors did not adjust for differential contributions of each twin pair to overall data throughout the study period</p>
Selective reporting (reporting bias)	High risk	<p>All outcomes that study authors set out to measure were reported, although the main outcome - parental self efficacy - was not measured. Study authors cited the reason as "Data were insufficient to analyse parental self-efficacy and parental stress" (Results, paragraph 1)</p> <p>Table 1 provides a matrix on outcomes listed in the methods, any additional outcomes that review authors considered as key outcomes for the study, and actual outcomes reported</p>
Other bias	High risk	<p>Further correspondence with the lead study author revealed the ways that major outcomes were reported and their associated risks of bias</p> <p>For the outcomes of "time in quiet sleep," "time crying," and "co-regulated," the unit of measurement was the number of times infants were observed to be in quiet sleep, crying, or co-regulated during each period of observation that lasted for 5 hours. A total of 8 five-hour blocks of observational periods were allocated over the study duration of 2 weeks. For each outcome, the number of episodes observed for each infant throughout the 2-week study duration was analyzed. Study authors reported mean numbers of episodes (with standard deviations and standard errors) per infant per 5-hour block of observation for co-bedded group vs control group</p> <p>From the way the data were analyzed and reported, as detailed above, the study appeared to adopt a repeated-measures design. Study authors reported overall mean numbers of episodes of each outcome for intervention and control groups; this was an acceptable way of reporting (Higgins 2011). However, 2 concerns arose: Study authors reported the data as per each infant instead of per each twin pair, giving rise to unit of analysis issues; and in the final estimates, no differential weight was given to data collected from each twin pair, as not all twin pairs were observed throughout the 2-week period (see Incomplete outcome data (attrition bias))</p>
Blinding of participants and personnel (performance bias) All outcomes	High risk	Although not stated in the article, blinding of caregivers and parents appeared highly unlikely

Hayward 2007 (Continued)

Blinding of outcome assessment (detection bias) All outcomes	High risk	Although not stated in the article, further correspondence with the lead author confirmed that the research assistant who collected the data was not blinded to the status of the twin pairs
---	-----------	--

Lutes 2001

Methods	Randomized controlled trial at a tertiary-level NICU (USA)
Participants	62 co-bedded infants (23 pairs of twins, 4 set of triplets, and 1 set of quadruplets) were compared with 59 separately bedded infants (15 pairs of twins, 7 sets of triplets, and 2 sets of quadruplets). All infants were less than 37 weeks' PMA. Birth weight was supposed to be less than 1500 grams, but in the study sample description (Table 4), the birth weight range went up to 2290 g. All infants were non-ventilated and were not receiving any other forms of intensive care; all had no congenital malformations nor severe neurosensory defects
Interventions	Co-bedding (cared for in the same incubator) vs separate care (cared for in separate incubators)
Outcomes	Average weekly gains in weight, head circumference, and length; feeding advancement, medication errors, infections, thermal insults
Notes	<ul style="list-style-type: none"> Results reported were combined results of twins and higher-order multiples, with no subgroup analysis for twins. We have contacted study authors to request raw data from which data for their twins can be extracted. We are still awaiting their reply Numbers analyzed, hence units of analysis for major outcomes of weight gain, head circumference, and length gains, were not stated. However, for outcomes of clinical issues (medication errors, nosocomial infections, sepsis workups initiated, thermal insults), the numbers of participants given (56 in the experimental group, 60 in the control group) suggest that analysis was based on individual infants, not on sets of twins or higher-order multiples

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Study authors stated, "A randomized design was used, with experimental (co-bedded) and control (traditional) conditions," but no other statement was provided on how randomization was achieved. Study authors also stated that "a convenient, consecutive sampling strategy was used to screen all twins and HOM (higher order multiples) according to the inclusion and exclusion criteria"
Allocation concealment (selection bias)	Unclear risk	No statement indicated how allocation was carried out
Incomplete outcome data (attrition bias) Sepsis	High risk	For the main outcome of growth, no mention was made of how many pairs of twins were actually analyzed and reasons for loss of data, if any. For outcomes of medication errors, nosocomial infections, sepsis workups initiated, and thermal insults, the number of infants analyzed was reported, and no data were lost
Selective reporting (reporting bias)	High risk	Although all outcomes specified in the methods were reported in the results with appropriate units of measurement, standard deviations for outcomes of weekly growth in weight, head circumference, and length, as well as total numbers of participants (sets of twins and higher-order multiples) for these outcomes were not reported Table 1 provides a matrix on outcomes listed in the methods, additional key outcomes, and actual outcomes reported

Lutes 2001 (Continued)

Other bias	Low risk	None identified
Blinding of participants and personnel (performance bias) All outcomes	High risk	Although not stated in the article, blinding of caregivers and parents appeared highly unlikely
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	It was not stated whether the research assistant who collected the data was blinded to the allocation of twins

NCASAR: Nursing Child Assessment Sleep/Activity Record

NICU: neonatal intensive care unit

NIDCAP: Newborn Individualized Developmental Care and Assessment Program

PDA: patent ductus arteriosus

PIPP: Premature Infant Pain Profile

PMA: postmenstrual age

RCT: randomized controlled trial

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Anon 2002	Commentary on another study, Stainton 2005 . Excluded on the basis of article type
DellaPorta 1998	Commentary that described development of a guideline on co-bedding of twins in a NICU with no original study. Excluded on the basis of article type
Fischer 1991	Study of co-sleeping between mother and baby, not co-bedding for twins. Excluded on the basis of the research question
Jahanfar 2012	Cross-sectional study on parents of twins 18 years of age and younger on the sleeping pattern and developmental milestones of the twins, as well as parents' knowledge and satisfaction with co-bedding. Excluded on the basis of study design
LaMar 2006	Retrospective descriptive study examining the incidence of infection in preterm twins over 2 periods: 1997 to 2001, when co-bedding was introduced as standard practice for preterm twins in the unit; and 1992 to 1996, before co-bedding was introduced. Excluded on the basis of study design
Longobucco 2002	Prospective cohort study comparing co-bedded multiples (twins and triplets) (31 infants in total) with historical controls comprising 31 infants of multiple gestations who were matched for PMA and birth weight. Major outcomes assessed were number of episodes of body temperature depression, daily average weight gain, and other physiological parameters. This study was published as an abstract. Excluded on the basis of study design
Lutes 2000	This record has no abstract, so no information on methods, participants, intervention, and outcomes. It appears very similar to one of the included studies (Lutes 2001), although we could not confirm this, as we have received no reply from study authors to multiple emails sent. Non-response of study authors led to exclusion of this article on the basis of lack of basic information
Matthews 2004	Prospective case-control study on sudden infant death syndrome (SIDS), not on co-bedding for preterm twins. Excluded on the basis of the research question
Mazeiras 2010	Review article describing history and current evidence on benefits and risks of co-bedding preterm twins. Excluded on the basis of article type

Study	Reason for exclusion
Nyqvist 1998	Single-group study in which 7 pairs of preterm twins were co-bedded. Data were collected on parental perception of infant behavior and on care before and after the twins were co-bedded. Excluded on the basis of study design
Nyqvist 2002	Prospective study examining breast-feeding behavior and intake among preterm twins and factors in caregiving practices that might influence them, including co-bedding of twins. This was not a study involving co-bedding as an intervention. Excluded on the basis of the research question
Polizzi 2003	Retrospective cohort study comparing co-bedded and separately bedded twins (n = 71 pairs) and triplets (n = 3 sets), with discharge weight, number of days on a ventilator, other clinical complications, and proportions co-bedded after discharge as major outcomes. Excluded on the basis of study design
Stainton 2005	Prospective cross-over trial in which 2 groups of twins - 1 group co-bedded for 24 hours or longer, and the other group with no prior co-bedding experience - were observed via video recording for 5 consecutive 30-minute periods, during which each group was co-bedded and separated for 30 minutes on an alternate basis. Excluded on the basis of the intervention, as we considered that alternation between co-bedded and separated states every 30 minutes would introduce an element of acute change in the infant environment, in addition to effects of co-bedding, which in turn would affect infants' neurobehavioral state. Additionally, it was highly likely that estimates at each state would have been "contaminated" by effects of the intervention during the preceding 30-minute period
Touch 2002	Single-group study in which all 11 pairs of twins were co-bedded. Data collected included apnea, periodic breathing, bradycardia, and other physiological parameters. Excluded on the basis of study design
Wahl 2006	Commentary on maternal-infant co-sleeping and breast-feeding in Norwegian, not an original study on co-bedding of preterm twins. Excluded on the basis of article type

NICU: neonatal intensive care unit

PMA: postmenstrual age

Characteristics of ongoing studies [ordered by study ID]

NCT01480856

Trial name or title	Co-bedding in Daily Weight Gain of Neonate Twins
Methods	RCT
Participants	<p>Inclusion criteria</p> <ul style="list-style-type: none"> At time of inclusion, twins must have been separated since less than 96 hours after birth Twins born between 30 and 34 weeks' gestation No severe congenital pathology Hospitalized at NICU Parents must have given non-opposition to this study within 96 hours after birth of twins. Time length of co-bedding estimated by the investigator to be 3 weeks <p>Exclusion criteria</p> <ul style="list-style-type: none"> Inclusion criteria not fulfilled Safety reasons Prolonged lack of comfort clinically harmful for the newborn, as per investigator conclusion

NCT01480856 (Continued)

Interventions	Co-bedding
Outcomes	Primary outcome: daily weight gain of co-bedded preterm twins vs that of single-bedded preterm twins Secondary outcomes: cardiorespiratory stability and other physiological measurements, newborn comfort, length of hospitalization, neuromotor development estimated by "Brunet Lezine" test at 2 years of age
Starting date	September 2008
Contact information	Nantes University Hospital (investigator details not provided)
Notes	

NICU: neonatal intensive care unit
RCT: randomized controlled trial

DATA AND ANALYSES

Comparison 1. Co-bedding vs standard care

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Rate of weight gain (in gram/kg of baseline weight/d)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.1 From study entry to week 1	1	38	Mean Difference (IV, Fixed, 95% CI)	4.0 [-0.96, 8.96]
1.2 From week 1 to week 2	1	28	Mean Difference (IV, Fixed, 95% CI)	1.40 [-2.27, 5.07]
1.3 From week 2 to week 3	1	18	Mean Difference (IV, Fixed, 95% CI)	-2.10 [-4.33, 0.13]
1.4 Average from study entry to week 3	1	18	Mean Difference (IV, Fixed, 95% CI)	0.20 [-1.60, 2.00]
2 Apnea, bradycardia, or desaturation	1	124	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.18, 4.05]
3 Episodes in co-regulated state (out of 20 observations)	1	6	Mean Difference (IV, Fixed, 95% CI)	0.96 [-3.44, 5.36]
4 Episodes of crying (out of 20 observations)	1	6	Mean Difference (IV, Fixed, 95% CI)	4.43 [1.72, 7.14]
5 Episodes in quiet sleep (out of 20 observations)	1	6	Mean Difference (IV, Fixed, 95% CI)	4.58 [1.58, 7.58]
6 Neurobehavior: infant pain score following painful procedure	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 At 30 seconds post heel lance	2	224	Mean Difference (IV, Fixed, 95% CI)	-0.96 [-1.68, -0.23]
6.2 At 90 seconds post heel lance	1	124	Mean Difference (IV, Fixed, 95% CI)	1.0 [0.14, 1.86]
7 Infection	3		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
7.1 Suspected or proven infection (any)	3	65	Risk Ratio (M-H, Fixed, 95% CI)	0.84 [0.30, 2.31]
7.2 Necrotizing enterocolitis	1	41	Risk Ratio (M-H, Fixed, 95% CI)	1.90 [0.19, 19.40]
7.3 Conjunctivitis	1	41	Risk Ratio (M-H, Fixed, 95% CI)	0.95 [0.15, 6.13]
7.4 Sepsis	2	59	Risk Ratio (M-H, Fixed, 95% CI)	0.45 [0.07, 2.86]
8 Length of hospital stay (days)	1	6	Mean Difference (IV, Fixed, 95% CI)	-4.90 [-35.23, 25.43]
9 Parental anxiety (Parental State Anxiety Inventory)	1	18	Mean Difference (IV, Fixed, 95% CI)	0.90 [-2.13, 3.93]
10 Parental attachment (Maternal Attachment Inventory)	1	18	Mean Difference (IV, Fixed, 95% CI)	0.90 [-2.02, 3.82]
11 Parental satisfaction	1	18	Mean Difference (IV, Fixed, 95% CI)	-0.38 [-4.49, 3.73]

Analysis 1.1. Comparison 1 Co-bedding vs standard care, Outcome 1 Rate of weight gain (in gram/kg of baseline weight/d).

Analysis 1.2. Comparison 1 Co-bedding vs standard care, Outcome 2 Apnea, bradycardia, or desaturation.

Analysis 1.3. Comparison 1 Co-bedding vs standard care, Outcome 3 Episodes in co-regulated state (out of 20 observations).

Analysis 1.4. Comparison 1 Co-bedding vs standard care, Outcome 4 Episodes of crying (out of 20 observations).

Analysis 1.5. Comparison 1 Co-bedding vs standard care, Outcome 5 Episodes in quiet sleep (out of 20 observations).

Analysis 1.6. Comparison 1 Co-bedding vs standard care, Outcome 6 Neurobehavior: infant pain score following painful procedure.

Analysis 1.7. Comparison 1 Co-bedding vs standard care, Outcome 7 Infection.

Analysis 1.8. Comparison 1 Co-bedding vs standard care, Outcome 8 Length of hospital stay (days).

Analysis 1.9. Comparison 1 Co-bedding vs standard care, Outcome 9 Parental anxiety (Parental State Anxiety Inventory).

Analysis 1.10. Comparison 1 Co-bedding vs standard care, Outcome 10 Parental attachment (Maternal Attachment Inventory).

Analysis 1.11. Comparison 1 Co-bedding vs standard care, Outcome 11 Parental satisfaction.

ADDITIONAL TABLES

Table 1. Matrix on outcome reporting by each study

Study ID	Outcomes listed in the methods	Additional key out-comes relevant to the study	Actual outcomes reported
Byers 2003a	<ul style="list-style-type: none"> Physiological measures including heart rate, respiratory rate, and oxygen saturation Sleep-wake synchrony Neurobehavioral observations using Newborn Individualized Developmental Care and Assessment Program (NIDCAP) guidelines Complications such as infection, intraventricular hemorrhage, and patent ductus arteriosus Parental measures including anxiety, attachment, and satisfaction 	<p>This study assessed specifically physiological measures in the short term. Growth parameters were reported, but as "mean daily weight," which was not suitable for meta-analysis. Additionally, physiological measures were reported in the form of average figures, such as average "highest activity heart rate," not as episodes of apnea, bradycardia, or desaturation, which are more clinically relevant</p>	<ul style="list-style-type: none"> Physiological measures including heart rate, respiratory rate, and oxygen saturation Sleep-wake synchrony was reported narratively; only ranges were given in the article Neurobehavioral observations based on Newborn Individualized Developmental Care and Assessment Program (NIDCAP) guidelines Complications such as caregiver error, intraventricular hemorrhage, and patent ductus arteriosus Parental measures including anxiety, attachment, and satisfaction Mean daily weight

Table 1. Matrix on outcome reporting by each study (Continued)

Campbell-Yeo 2012b	<ul style="list-style-type: none"> Pain response measured by the Premature Infant Pain Profile (PIPP) Time for physiological recovery in response to heel lance determined by length of time for heart rate and oxygen saturation to return to baseline Need for additional pain relief with 24% sucrose Adverse outcomes such as episodes of apnea, bradycardia, infection, and caregiver error 	<p>This study assessed specifically infants' response to pain; growth-related outcomes such as weight gain and length of hospital stay were not included in the outcomes. The only outcome that was relevant to this review was the pain response, which we considered as a form of neurobehavior</p>	<ul style="list-style-type: none"> Pain response measured by the Premature Infant Pain Profile (PIPP) Time for physiological recovery in response to heel lance determined by length of time for heart rate and oxygen saturation to return to baseline Need for additional pain relief with 24% sucrose Adverse outcomes such as episodes of apnea, bradycardia, infection, and caregiver error
Chin 2006	<ul style="list-style-type: none"> Weight gain Apnea/bradycardia/desaturation (A/B/D) episodes Infection, including sepsis (positive blood culture), necrotizing enterocolitis (clinical and radiological diagnosis), or conjunctivitis (positive eye culture) 	<ul style="list-style-type: none"> Rate of weight gain Length of hospital stay Medication error 	<ul style="list-style-type: none"> Weight A/B/D episodes Infection, including sepsis, necrotizing enterocolitis, or conjunctivitis Medication error
Hayward 2007	<ul style="list-style-type: none"> Parental self efficacy (measured using the Infant Care Survey) and parental anxiety (measured using the Spielberger State-Trait-Anxiety Inventory) upon entry to the study, before discharge, and at 1-month post discharge Co-regulatory behavior (measured using the Nursing Child Assessment Sleep/Activity Record, which included assessment of synchrony of infant state (e.g. in quiet sleep, alert, crying), heart rate, temperature, respirations, and oxygen saturation Infection rate ("incidence of septic workup, treatment with antibiotics, and confirmed incidence of sepsis") Caregiver error 	<ul style="list-style-type: none"> Length of hospital stay Weight gain Apnea, bradycardia, or desaturation episodes 	<ul style="list-style-type: none"> Time in quiet sleep and crying Time co-regulated Infection rate Caregiver error Length of hospital stay (as part of descriptive characteristics, not part of results) <p>(Study authors stated, "Data were insufficient to analyse parental self-efficacy and parental stress" - Results, paragraph 1, lines 1 to 3)</p>
Lutes 2001	<p>No outcome was listed in the methods. However, the following 3 major outcomes were stated in the purpose and hypothesis section</p> <ul style="list-style-type: none"> Weight gain Head circumference growth Longitudinal growth/length gain 	<ul style="list-style-type: none"> Length of hospital stay Apnea, bradycardia, or desaturation episodes Infection Medication error 	<ul style="list-style-type: none"> Weekly weight gain "Weekly mean kilocalories per kilogram" Weekly head circumference growth Weekly length growth Medication error Nosocomial infection Sepsis workups initiated Thermal insults
Badiee 2014	<ul style="list-style-type: none"> Pain response to heel prick measured by the Premature Infant Pain Profile (PIPP) Infant salivary cortisol level before and after heel prick procedure 	<p>Like Campbell-Yeo 2012b, this study assessed specifically infant response to pain;</p>	<p>The only outcome included in this review is PIPP score</p>

Table 1. Matrix on outcome reporting by each study (Continued)

growth-related out-
comes were not assessed

Table 2. Median combined apnea, bradycardia, and desaturation (A/B/D) episodes

Study ID	Study period	Median combined A/B/D episodes		P value
		Co-bedded group	Control group	
Chin 2006	Week 1	4.5	7	0.2
	Week 2	6	12	0.8
	Week 3	2.5	8	0.4

APPENDICES

Appendix 1. Standard search strategy

PubMed: ((infant, newborn[MeSH] OR newborn OR neonate OR neonatal OR premature OR low birth weight OR VLBW OR LBW or infan* or neonat*) AND (randomised controlled trial [pt] OR controlled clinical trial [pt] OR randomised [tiab] OR placebo [tiab] OR drug therapy [sh] OR randomly [tiab] OR trial [tiab] OR groups [tiab]) NOT (animals [mh] NOT humans [mh]))

EMBASE: (infant, newborn or newborn or neonate or neonatal or premature or very low birth weight or low birth weight or VLBW or LBW or Newborn or infan* or neonat*) AND (human not animal) AND (randomised controlled trial or controlled clinical trial or randomised or placebo or clinical trials as topic or randomly or trial or clinical trial)

CINAHL: (infant, newborn OR newborn OR neonate OR neonatal OR premature OR low birth weight OR VLBW OR LBW or Newborn or infan* or neonat*) AND (randomised controlled trial OR controlled clinical trial OR randomised OR placebo OR clinical trials as topic OR randomly OR trial OR PT clinical trial)

The Cochrane Library: (infant or newborn or neonate or neonatal or premature or preterm or very low birth weight or low birth weight or VLBW or LBW)

WHAT'S NEW

Date	Event	Description
6 February 2017	Amended	Added external source of support

HISTORY

Protocol first published: Issue 1, 2010

Review first published: Issue 12, 2012

Date	Event	Description
29 February 2016	New citation required but conclusions have not changed	Updated search performed to February 2016. One further study added to the meta-analysis

Date	Event	Description
		One study that was originally awaiting classification has now been excluded as the result of non-response of study authors to our repeated enquiries for information needed to determine eligibility In the "Summary of findings" table, we removed the outcome "PIPP score 90 seconds after heel lance" and replaced it with "PIPP score 30 seconds after heel lance" following inclusion of a new study with the latter outcome
29 February 2016	New search has been performed	This updates the review, "Co-bedding in neonatal nursery for promoting growth and neurodevelopment in stable preterm twins," published in <i>The Cochrane Library</i> (Lai 2012)
1 June 2011	Amended	Converted from protocol to full review

CONTRIBUTIONS OF AUTHORS

NML, SCF, WCF, and KT drafted the protocol.

NML, SCF, and WCF performed the search and identified relevant articles.

NML and SCF acquired full texts for articles identified to be potentially suitable.

NML and SCF independently assessed the eligibility of identified articles.

NML, SCF, WCF, and KT wrote the description of studies, the results, the discussion, and the conclusions.

NML wrote the abstract.

NML and SCF wrote the plain language summary.

All review authors approved the final draft.

NML updated the review with edits from the other review authors.

DECLARATIONS OF INTEREST

None declared by the review authors.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- SEA-ORCHID (South East Asian Optimising Reproductive and Child Health Outcomes for Developing Countries) Project, Malaysia.

A five-year (2003-2008) project, funded by the National Medical Research Council of Australia and Wellcome Trust, and supported by the Cochrane Australasian Centre, involving four South East Asian countries (Indonesia, Malaysia, Philippines, and Thailand) aiming to promote the synthesis and use of high-level clinical evidence particularly for topics in maternal and child health that are relevant to this part of the world

- Eunice Kennedy Shriver National Institute of Child Health and Human Development National Institutes of Health, Department of Health and Human Services, USA.

The Cochrane Neonatal Review Group has been funded in part with Federal funds from the Eunice Kennedy Shriver National Institute of Child Health and Human Development National Institutes of Health, Department of Health and Human Services, USA, under Contract No. HHSN267200603418C

- National Institute for Health Research, UK.

Editorial support for Cochrane Neonatal has been funded with funds from a UK National Institute of Health Research Grant (NIHR) Cochrane Programme Grant (13/89/12). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR, or the UK Department of Health.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We revised the following outcomes after receiving feedback from the external referee on our draft review.

- Primary outcome: neurobehavior - we have added "infant co-regulation" and "stress response" to our statement.
- Secondary outcome: adverse effects - we have added "and other patient safety-related outcomes as defined by the individual studies."
- A new outcome added: secondary outcome: incidence of co-bedding post discharge .

INDEX TERMS

Medical Subject Headings (MeSH)

*Beds; *Twins; Child Development [*physiology]; Infant Care [*methods]; Infant, Premature [*growth & development]; Intensive Care Units, Neonatal; Randomized Controlled Trials as Topic

MeSH check words

Humans; Infant, Newborn