Non-pharmacological management of infant and young child procedural pain (Review)

Pillai Riddell RR, Racine NM, Turcotte K, Uman LS, Horton RE, Din Osmun L, Ahola Kohut S, Hillgrove Stuart J, Stevens B, Gerwitz-Stern A

THE COCHRANE COLLABORATION®

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2011, Issue 10

http://www.thecochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Non-pharmacological management of infant and young child procedural pain

Rebecca R Pillai Riddell[1], Nicole M Racine[1], Kara Turcotte[1], Lindsay S Uman[2], Rachel E Horton[1], Laila Din Osmun[1], Sara Ahola Kohut[1], Jessica Hillgrove Stuart[1], Bonnie Stevens[3], Alanna Gerwitz-Stern[1]

[1]Department of Psychology, York University, Toronto, Canada. [2]IWK Health Centre & Dalhousie University, Halifax, Canada. [3]Associate Chief of Nursing Research, The Hospital for Sick Children, Toronto, Canada

Contact address: Rebecca R Pillai Riddell, Department of Psychology, York University, 4700 Keele Street, OUCH Laboratory, Atkinson College, Toronto, Ontario, M3J 1P3, Canada. rpr@yorku.ca. www.yorku.ca/ouchlab.

Editorial group: Cochrane Pain, Palliative and Supportive Care Group.
Publication status and date: Edited (no change to conclusions), published in Issue 5, 2012.
Review content assessed as up-to-date: 25 July 2011.

Citation: Pillai Riddell RR, Racine NM, Turcotte K, Uman LS, Horton RE, Din Osmun L, Ahola Kohut S, Hillgrove Stuart J, Stevens B, Gerwitz-Stern A. Non-pharmacological management of infant and young child procedural pain. *Cochrane Database of Systematic Reviews* 2011, Issue 10. Art. No.: CD006275. DOI: 10.1002/14651858.CD006275.pub2.

ABSTRACT

Background

Infant acute pain and distress is commonplace. Infancy is a period of exponential development. Unrelieved pain and distress can have implications across the lifespan.

Objectives

To assess the efficacy of non-pharmacological interventions for infant and child (up to three years) acute pain, excluding breastmilk, sucrose, and music. Analyses accounted for infant age (preterm, neonate, older) and pain response (pain reactivity, pain-related regulation).

Search methods

We searched CENTRAL in *The Cochrane Library* (2011, Issue 1), MEDLINE (1966 to April 2011), EMBASE (1980 to April 2011), PsycINFO (1967 to April 2011), Cumulative Index to Nursing and Allied Health Literature (1982 to 2011), Dissertation Abstracts International (1980 to 2011) and www.clinicaltrials.gov. We also searched reference lists and contacted researchers via electronic listserves.

Selection criteria

Participants included infants from birth to three years. Only randomized controlled trials (RCTs) or RCT cross-overs that had a no-treatment control comparison were eligible for inclusion in the analyses. We examined studies that met all inclusion criteria except for study design (e.g. had an active control) to qualitatively contextualize results.

Data collection and analysis

We refined search strategies with three Cochrane-affiliated librarians. At least two review authors extracted and rated 51 articles. Study quality ratings were based on a scale by Yates and colleagues. We analyzed the standardized mean difference (SMD) using the generic inverse variance method. We also provided qualitative descriptions of 20 relevant but excluded studies.

Main results

Fifty-one studies, with 3396 participants, were analyzed. The most commonly studied acute procedures were heel-sticks (29 studies) and needles (n = 10 studies). The largest SMD for treatment improvement over control conditions on pain reactivity were: non-nutritive sucking-related interventions (preterm: SMD -0.42; 95% CI -0.68 to -0.15; neonate: SMD -1.45, 95% CI -2.34 to -0.57), kangaroo care (preterm: SMD -1.12, 95% CI -2.04 to -0.21), and swaddling/facilitated tucking (preterm: SMD -0.97; 95% CI -1.63 to -0.31). For immediate pain-related regulation, the largest SMDs were: non-nutritive sucking-related interventions (preterm: SMD -0.38; 95% CI -0.59 to -0.17; neonate: SMD -0.90, 95% CI -1.54 to -0.25), kangaroo care (SMD -0.77, 95% CI -1.50 to -0.03), swaddling/facilitated tucking (preterm: SMD -0.75; 95% CI -1.14 to -0.36), and rocking/holding (neonate: SMD -0.75; 95% CI -1.20 to -0.30). The presence of significant heterogeneity limited our confidence in the lack of findings for certain analyses.

Authors' conclusions

There is evidence that different non-pharmacological interventions can be used with preterms, neonates, and older infants to significantly manage pain behaviors associated with acutely painful procedures.

PLAIN LANGUAGE SUMMARY

Non-pharmacological interventions for acute pain in infants

We examined 13 different types of commonly investigated non-pharmacological treatments (excluding breastmilk, sucrose, and music) to determine their efficacy for pain reactions after an acutely painful procedure (right after the needle ('pain reactivity') and less immediate pain reactions ('immediate pain-related regulation'). Fifty-one randomized controlled trials were included involving 3396 participants. For preterm infants, there was sufficient evidence to recommend kangaroo care, sucking-related interventions, and swaddling/facilitated tucking interventions for both pain reactivity and immediate pain-related regulation. For neonates, there was sufficient evidence to recommend sucking-related interventions as an effective treatment for pain reactivity and immediate pain-related regulation. Rocking/holding was also found to be efficacious for neonatal immediate pain-related regulation. For older infants, there were no treatments reviewed that demonstrated sufficient evidence. Due to significant differences in the magnitude of treatment effects among studies (heterogeneity), some analyses that found a lack of treatment effect need to be interpreted with caution.

BACKGROUND

Description of the condition

Despite the vigorous responses that result when an infant is subjected to a painful procedure, the premise that infants are insensitive to pain was only recently rejected by the general scientific community (McClain 2005), although exceptions in the literature still remain (e.g. Derbyshire 1999). Early studies suggested that infants did not possess a cortex well developed enough to perceive or localize pain (McGraw 1943). Moreover, initial misinterpretations of common infant pain outcomes, for example, the lack of declarative memory for painful experiences during infancy (Field 1995), the muted responses of premature infants after a barrage of painful procedures (Johnston 1993), and unacceptable rates of cardiac arrest or death due to poor knowledge of infant morphokinetics during the 1950s to 1970s (Berde 2005), all perpetuated widespread neglect of infant pain treatment until the last three decades.

Current research supports the understanding that infants possess the anatomical and functional requirements to perceive pain (Fitzgerald 2005; Slater 2010) and respond during tissue insult in a manner unequivocally interpretable as pain (Grunau 1987). However, despite evidence of the long-term implications of unrelieved pain during infancy (Anand 2000; Grunau 1996; Grunau 2000; Howard 2003; Taddio 1997), evidence that infant pain is still undermanaged and unmanaged is clearly evident (Alexander 2003; Pillai Riddell 2005; Simons 2003; Taddio 2010). A comprehensive and systematic review of pain management strategies will be integral to appropriate infant pain management.

Description of the intervention

Generally speaking, pain management can be subdivided into two categories: pharmacological (Barber 2004; treatments that deal with the uses, effects, and modes of action of drugs) and non-pharmacological (any treatments (contextual, psychological, and behavioral strategies) that do not deal with uses, effects, and modes of action of drugs).

In a recent précis of infant non-pharmacological pain management strategies, pain management was defined as any strategy or technique administered to an infant in pain with the intention of lessening pain sensation, pain perception, or both (Pillai Riddell 2006). One of the important principles in infant pain management is to recognize that pain is most effectively managed by avoiding, preventing or limiting exposure to pain-provoking stimuli (Joint Committee 2000). Accordingly, pain management during infancy must be multifaceted and integrated within every step of the decision-making process; from deciding whether a particular procedure is warranted to determining the safest and most efficacious pain relieving strategy.

There are comprehensive reviews which summarize assessment and management techniques for painful procedures in neonates or infants, or both, which have recently been published (Anand 2001; Cignacco 2007; Cohen 2008; Johnston 2011; Kleiber 1999; Obeidat 2009; Piira 2005; Prasopkittikun 2003; Schechter 2007; Slifer 2002). However, to our knowledge, there has been no comprehensive, systematic meta-analysis conducted on the efficacy or effectiveness of non-pharmacological interventions for managing acute pain and distress in infants and young children up to three years of age that:

1. analyze results according to developmental subcategories within infancy (i.e. infant born preterm, healthy neonate, older infant); and
2. analyze type of pain response (i.e. responses right after the painful stimulus (pain reactivity) versus those after the initial pain response period (pain-related regulation)).

Why it is important to do this review

Both dimensions of analysis are crucial due to the steep trajectory of infant development, both psychologically and physiologically. Moreover, given the different physiological and psychological mechanisms subsuming a person's immediate reaction to a painful stimulus (more automatic and/or reflexive reactivity) and a person's reaction during the period of recovering from the painful insult (pain-related regulation), it was seen as crucial to elucidate this variable to move the infant acute pain management literature forward. Therefore, this review will take a broader and more in-depth look at the non-pharmacological pain management literature for infant acute pain.

OBJECTIVES

To assess the efficacy of non-pharmacological interventions for the management of acute pain and distress in infants and children up to three years of age, with specific attention to preterm births, healthy neonates, and older infants. Analyses will include separate procedures for pain reactivity right after the acute pain stimuli and pain-related regulation responses that occur after the initial period of pain reactivity.

METHODS

Criteria for considering studies for this review

Types of studies

We only included randomized controlled trials (RCTs) and randomized cross-over trials that utilized a no-treatment control group and involved the management of acute procedural pain in infants and children up to three years of age in this review. We used no language restrictions during the search.

Due to existing work completed by other Cochrane Review authors, this review excluded studies that focused on sucrose (Stevens 2010), breast milk (Shah 2009), or music (Cepeda 2006) as a pain management strategy. In addition, due to other existing Cochrane Reviews on these acute pain stimuli, we also excluded studies that examined pain management for the following types of acute pain stimuli and age group combinations: circumcision procedure for boys aged zero to three years (Brady-Fryer 2009; Cyna 2010), blood sampling via heel lance or venipuncture in neonates up to 28 days (Shah 2009), and needle-related or procedural pain in children older than three years (Uman 2006).

Types of participants

Participants included all young children who were undergoing painful acute procedures. Term and preterm infants were also included in the review. Given that research in the area of infant pain management began in the late 1980s, we selected a broad mandate of 'procedural pain' rather than any particular type of procedure. However, in order to provide general parameters regarding procedures that are under the review, sample procedures are provided. Based on two comprehensive references that outline painful procedures in either neonates or older children (Anand 2001; Uman 2006), the following non-exhaustive list is provided as a sample of procedures that fell under the umbrella of this review (*see* 'Table 1'). Definitions were derived from two online medical encyclopedic reference sources (i.e., MEDLINE Plus Medical Encyclopedia: www.nlm.nih.gov/medlineplus/mplusdictionary.html; the Merck Manual of Diagnosis and Therapy, 17th Edition, www.merck.com) and by consulting with medical professionals in the area of infant pain.

Types of interventions

We clustered non-pharmacological interventions into three different categories based on their hypothesized mechanism of action (Pillai Riddell 2006):

Contextual strategies

The context in which a painful procedure is conducted modifies behavioral and physiological expression of infant pain. Interventions that are classified in this category involved modifying the environment to have lower pain reactivity and stress. They tended to involve multiple components that in combination modify the environment in which an infant experiences pain (i.e. low noise and lighting, clustering procedures to avoid over handling, soothing smells).

Cognitive strategies

Any intervention that is suspected to have a mechanism of action that impacts an infant's abilities to perceive the pain experience was classified here. The main intervention falling under this category is distraction. We defined distraction as any procedure aimed at diverting infant attention from the painful stimulus. For the purposes of the review, we defined distraction as toy-mediated distraction (use of a children's toy) versus video-mediated distraction (use of an audio-visual screen displaying two-dimensional moving images with co-ordinated audio output).

Behavioral strategies

These strategies involve either direct (e.g. rocking) or indirect (e.g. non-nutritive sucking; care-giver provides soother) manipulation of the infant's body by a care-giver. Most research on non-pharmacological pain management strategies has been conducted within this domain. Accordingly, a number of strategies are covered in this review. For ease of consumption, we grouped strategies that were considered to have a similar (albeit not identical) mechanism of action together into categories. The categories of treatments that were included in the statistical analysis were the following.

1. Kangaroo care (also known as skin-to-skin contact): an infant is placed on their care-giver's bare chest during a painful procedure or for soothing after a painful procedure.
2. Swaddling/facilitated tucking: swaddling is when an infant is securely wrapped in a blanket to prevent the child's limbs from moving around excessively. Facilitated tucking involves firmly containing the infant using a care-giver's hands on both head and lower limbs to maintain a 'folded-in' position. Infant may or may not be wearing clothes.
3. Non-nutritive sucking-related strategies: an object (e.g. pacifier, non-lactating nipple) is placed into an infant's mouth to stimulate oro-tactile or sucking behaviors during a painful event. This may have involved other adjuvant non-pharmacological interventions that fall under the purvey of the review (e.g. pacifier plus water was included; pacifier plus sucrose was not included).
4. Swallowing water: water is administered for ingestion without an instrument that would incite extensive sucking (e.g. water administered by a dropper).
5. Rocking, holding or both: an infant is held or gently moved up and down or side to side (or both) by a care-giver.
6. Simulated rocking and water: as opposed to being held by an adult, an infant is placed in a bassinet-type machine that provides a swaying motion. In addition, water was administered in a manner that did not incite extensive sucking.
7. Touch/massage: an infant's body is 'stroked' to provide some type of counter-stimulation to the nociceptive input.
8. Structured parental involvement: parents are instructed or informed of strategies that are accepted as pain-reducing but are not given any materials to aid them. A variety of strategies may or may not be enacted such as rocking, holding, shushing, talking, rubbing, tickling, and distracting attention without toy or video.
9. Maternal voice: an infant is exposed to a reproduction of his or her mother's voice, designed to help simulate the fetal environment.
10. Parent present: simply allowing the parent to be present during a painful procedure but parents are not interacting extensively with their child in a manner thought to be pain-reducing.

A number of strategies were found in the treatment literature which could not be included in the review because they did not have at least one study that met the criteria for the quantitative analysis, most often due to the choice of an active control group. These strategies were olfactory stimulation (providing either a pleasing smell or a familiar smell before, during and after a painful procedure), order of immunizations (providing the most painful immunization last), position during procedure (infant is positioned either prone or supine during procedure), and formula (providing infant formula during the painful procedure).

Types of outcome measures

Due to the limited verbal capacity of the infant it is important to recognize that pain measures are limited in distinguishing between infant pain and infant distress (Craig 2002). However, due to the presence of an objectively painful stimulus in all studies selected for this review, we considered all measures of negative reactions after the administration of a known painful stimulus an indicator of an infant's pain.

Since the purpose of this review was to be able to outline intervention-specific, age-specific, and pain response-specific information, outcomes were not subdivided by type of outcome measure. Due to the emerging verbal skills of infants and recommended clinical assessment procedures for infant pain (Franck 2000; Stevens 2007), we made the decision to only analyze objectively measured behavioral responses to pain. When studies had more than one

behavioral response to pain, we used the most specific measure available. For example, we used pain facial expression over cry duration. We kept detailed logs about decisions regarding which measure (when multiple were available) was selected from an article and why. The lead author reviewed all tables to confirm judgments made by other review authors.

Physiological measures are rarely used in clinical practice as the sole indicator of pain and the literature showed considerable inter-study variance with the measurement of physiological indicators, therefore they were not analyzed for the review. No article was completely excluded from the review because it only used physiological indicators. These articles were still qualitatively mentioned in the respective 'Summary of treatment effects' section (at the end of every treatment effect section description, results for each age group and pain response type are summarized briefly). Finally, if well-established, multi-dimensional pain measures were utilized, with behavioral and physiological indicators, and it was a reliable and valid measure, we used the total score (for example, the Premature Infant Pain Profile [PIPP; Stevens 1996]). The pain measurements used in the included studies were as follows:

- **Premature Infant Pain Profile (PIPP)** (13 studies; Akcan 2009; Axelin 2009; Bellieni 2001; de Sousa 2008; Elserafy 2009; Hill 2005; Johnston 2003; Johnston 2007; Kozub 2001; Liaw 2010, Sizun 2002; Stevens 1999; Ward-Larson 2004).
- **Duration of cry** (eight studies; Allen 1996; Bauchner 1996; Blass 1999; Campos 1994; Corff 1995; Greenberg 2002; Herrington 2007; Kostandy 2008).
- **Neonatal Facial Coding System (NFCS)** (eight studies; Bustos 2008; Castral 2008; Chermont 2009; Comaru 2009; Fearon 1997; Gormally 2001; Ipp 2004; Johnston 1997).
- **Neonatal Infant Pain Scale (NIPS)** (nine studies; Axelin 2006; Bo 2000; Catelin 2005; Im 2008; Jain 2006; Kashaninia 2008; Liu 2010, Morrow 2010; Yilmaz 2010).
- **Modified Behavioral Pain Scale (MBPS)** (four studies; Cohen 2002; Cramer-Berness 2005; Cramer-Berness 2005b; Hillgrove Stuart 2008).
- **Douleur Aiguë Nouveau-né (DAN)** (three studies; Bellieni 2002; Carbajal 1999; Carbajal 2003).
- **Newborn Individualized Developmental Care and Assessment Program (NIDCAP)** (one study; Ferber 2008).
- **Infant behavioral state** (one study; Whipple 2004).
- **Grimace** (one study; Gray 2000).
- **Face Legs Arms Cry Consolability Scale (FLACC)** (one study; Curtis 2007).
- **Measure of Adult and Infant Soothing and Distress (MAISD)** (one study; Cohen 2006).
- **Brazelton Neonatal Behavioral Assessment Scale (BNBAS)** (one study; Corbo 2000).

Search methods for identification of studies

Electronic searches

In terms of published studies, we designed a unique search strategy for each of four databases in conjunction with three librarians affiliated with The Cochrane Collaboration. They were MEDLINE (1966 to 2011), EMBASE (1980 to 2011), PsycINFO (1967 to 2011), and a Cumulative Index to Nursing and Allied Health Literature (1982 to 2011). See Appendix 1 for a sample of the search strategy for published articles. The search ended in April 2011.

Searching other resources

In addition, we located completed trials that were not yet published through Dissertation Abstracts International (1980 to April 2011), the Cochrane Central Register of Controlled Trials (CENTRAL) in *The Cochrane Library* (2011, Issue 1), and www.clinicaltrials.gov (6 January 2010).

Also, we made appeals to pediatric list-serves (Pain in Child Health [PICH], Pediatric Pain, American Psychological Association's Division 54 [Pediatric Psychology] list-serve) for unpublished trials relating to non-pharmacological pain management in infants. List-serves are electronic mailing lists that allow for widespread distribution of information to many Internet users.

Finally, we used the reference lists of recently published reviews and meta-analyses cited in the introduction to ensure completeness.

Data collection and analysis

Selection of studies

Four review authors (RPR, LU, AG, KT) independently screened titles and abstracts of studies from literature searches for inclusion in the review. Review authors were not blinded to authors, institutions, journals, or results. For all abstracts where relevance was questionable, the full article was read by the lead review author (RPR) who made the final decision regarding inclusion. Using the full articles, six review authors (RPR, NR, RH, LDO, SAK, JHS) reviewed articles for inclusion. For all articles where relevance was questionable, the lead review author (RPR) made the final decision.

Data extraction and management

Six review authors (RPR, NR, RH, LDO, SAK, JHS) conducted data extraction in three teams of two (RPR and NR; RH and JHS; SAK and LDO) using a data extraction form designed for this specific review. Each form also included a risk of bias questionnaire that was designed for this review (based on Yates 2005), that required the review author to make subjective judgments on the study quality. All included articles for extraction were divided up equally between the three teams. Both members of a team extracted every included article independently. Every extracted data

point (including risk of bias questions) for every article was then compared between both members of the team to ensure accuracy. When the two members of a team disagreed, we consulted the original article for pertinent information. In the case of disputes between the two team members, the lead author (RPR) would consult the article to make the final decision.

A research assistant compiled qualitative data from the studies into an electronic database in Excel 2007. A research assistant compiled quantitative data from the studies in RevMan 5.1 (RevMan 2011). A research assistant also compiled study quality data from the risk of bias questionnaire in an electronic database in SPSS 19.0. A sub-team of review authors (KT, NR) double-checked a random sampling of at least 25% of all data entered across all three programs.

Assessment of risk of bias in included studies

At least two review authors of the extraction team (RPR, NR, RH, JHS, LDO, SAK) scored every study included in the review for quality. We used the Quality of Study Design and Methods Scale (Yates 2005) as the basis to develop the study quality rating form for this review. Although this scale could be used for a broad array of research areas, this scale was specifically validated with non-pharmacological treatments (psychological treatments). The original scale requires authors to make judgments on scales ranging from 0 to 1 or 0 to 2 on eight multi-part questions. The maximal attainable score is 26, with higher scores indicating higher quality. We made three minor modifications to the Yates scale based on lack of relevance to the present review. Item four has a part that relates to treatment expectations (i.e. patient expectations) and was not scored due to age of patient (less than three years). Second, item six (in the original scale) was completely omitted due to the focus of the current review on brief procedural pain interventions. In the original scale this item related to the following up of the patient for at least six months post-treatment. Finally, item eight relating to having an adequately matched control group was made to be out of one rather than two due to the difficulty of having a no-treatment control group with a structurally similar component to that of the treatment group. Active control groups were excluded from our quantitative analyses thus making this level of equivalence generally unattainable. An active control group is a control group that receives a treatment. The maximum score on the revised Yates scale is 23 points with higher scores indicating higher quality. In addition, based on the recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), extractors were also required to respond to additional open-ended questions.

In addition, it was felt that an additional indicator was required to reflect treatment integrity. Especially given 21 of our studies were of a cross-over design, whereby the same group received both treatment and control interventions. We utilized a qualitative scale that contained the descriptors 'good', 'satisfactory', 'poor', and 'unknown'. Each of the two raters were independently required to judge the treatment integrity (using the qualitative scale) on five dimensions: Treatment Adherence, Treatment Exposure, Quality of Delivery, Participant Responsiveness and Program Differentiation (Higgins 2011). Then, they were required to review their scores across the five domains and give an overall judgment as to the integrity of the intervention (using the qualitative scale).

Scale questions

1. Are the inclusion and exclusion criteria clearly specified (two parts)?
    i) Sample criteria: zero to one;
    ii) Evidence that criteria was met: zero to one.
2. Is there evidence that CONSORT guidelines for reporting attrition have been followed (two parts)?
    i) Attrition: zero to two;
    ii) Rates of attrition zero to one.
3. Is there a good description of the sample (two parts)?
    i) Sample characteristics: zero to one;
    ii) Group equivalence: zero to one.
4. Adequate steps to minimize bias (three parts)?
    i) Randomization: zero to two;
    ii) Allocation bias: zero to one;
    iii) Measurement bias: zero to one.
5. Are the outcomes that have been chosen appropriate (three parts)?
    i) Justification of outcomes: zero to two;
    ii) Validity of outcomes for context: zero to two;
    iii) Reliability and sensitivity: zero to two.
6. Are the statistical analyses adequate (five parts)?
    i) Power calculation: zero to one;
    ii) Sufficient sample: zero to one;
    iii) Planned analysis: zero to one;
    iv) Statistics reporting zero to one;
    v) Intention-to-treat: zero to one.
7. Has a good, well-matched alternative treatment group been used (one part)?
    i) Control group: zero to one.

Additional open-ended questions were also posed:

i) Any sort of blinding of personnel, participants, and outcome assessors/coders? (Yes or No - explain).
ii) Any other selective outcome reporting?
iii) Any other potential sources of bias? (e.g. baseline inequality of treatment groups on relevant measures).

Measures of treatment effect

It was originally proposed that analyses would also take into account pain measure utilized and pain source. However, there were insufficient data to conduct meaningful analyses with these dimensions. Thus, we first organized results according to treatment,

then age, then pain response type. For example, an analysis that was conducted was to determine if kangaroo care was effective for preterm infants' pain reactivity.

We treated all of the outcome data for the included studies as continuous. Due to the decision to include randomized control trials (RCTs) and randomized cross-over trials, we used different procedures, obtained from a Cochrane statistical consultant, to calculate the standardized mean difference (SMD) and the standard error of the mean difference (SE), depending on whether the study was parallel-group or cross-over. We calculated the standard error of the SMD as follows for parallel trials: the square root of: (n1 + n2/n1*n2 + [SMD*SMD/2*(n1+n2)]). The standard error of the SMD for cross-over trials used the square root of: 2*(1-r)*(1/n+[SMD*SMD/2n]). If a trial provided two arms for one treatment analysis, we divided the control group *n* such that they were not double-counted within the same analysis.

Dealing with missing data

For every study that met the inclusion criteria, we contacted study authors when data were missing. Finally, whenever possible, using the recommended techniques from the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), we interpolated/estimated data from P-values, t-scores, and F-scores. At least two review authors of the extraction team (RPR, NR, RH, JHS, LDO, SAK) carried out every calculation by hand to limit calculation or entry errors. As an additional level of quality assurance for the review, whenever data was interpolated from a study, a sub-team of authors (RPR, NR) also conducted a conceptual double-check ensuring confidence intervals (CIs) from interpolated data were always in line with the findings from the original authors.

Assessment of reporting biases

If more than 20% of the originally randomized participants were not available for the outcome analysis, the data were not to be incorporated in the statistical analysis. However, this event did not occur in any of the studies that were considered for this review.

Due to the limited number of studies in each comparison of treatment (13 different treatment types were included in the review) X age group (preterm, neonate, older infant) X pain type, funnel plots were not appropriate. However, in order to help overcome publication bias, we imposed no language barriers, contacted known infant pain management researchers through personal emails and list-serves, and utilized both dissertation and trial registration sites.

Data synthesis

Guiding principles

Two main principles that guided the meta-analysis of the data collected for this review:

1. Under the 13 different categories of treatments, we separated studies into one of three age subcategories:
   i) Preterm - infants born at 36 weeks gestation or less.
   ii) Neonate - infants born at 37 weeks until one month of age.
   iii) Older infant - infants over one month to 36 months of age.
2. Each study within an age subcategory was further divided into one of two pain response subcategories to reduce heterogeneity among included studies:
   i) Pain reactivity- infant pain response was measured within the first 30 seconds after the acutely painful stimulus was discontinued.
   ii) Pain-related regulation- infant pain response was measured after the first 30 seconds after the acutely painful stimulus. If multiple measurements were taken after the first 30 seconds elapsed, we utilized the measurement closest to the 30 second time point.

We pooled the results from individual treatment studies using the generic inverse method for a random-effects model in Review Manager 5.1 (RevMan 2011). Using this statistical methodology, an index of the variability of the sample (standard error) and the number of participants in the sample (sample size) are used to determine how influential each study will be in the final meta-analytic statistic. The greater the variability (generally associated with small sample sizes), the less a particular study would be weighted in the final analysis. In addition to the SMD, we also reported a 95% confidence interval (CI); which incorporates the standard error of the pooled treatment effect) for the treatment effect.

Subgroup analysis and investigation of heterogeneity

Although we utilized a statistical analysis plan that goes beyond typical reviews to date to limit heterogeneity among studies (i.e. treatment X age X pain response analyses), the existence of heterogeneity between studies was inevitable. Given our primary interest in the impact of the heterogeneity (not the presence of heterogeneity), we utilized the I^2 statistic (as cited in Deeks 2005). In cases where substantial heterogeneity was found, when possible (i.e. more than two studies were analyzed), we re-ran the analyses without the source of heterogeneity (noted using the SMD) and compared to the original findings. These data were included in the individual result summaries under each treatment. Insufficient data precluded formal quantitative analyses for potential reasons for heterogeneity.

Sensitivity analysis

We investigated factors that may have affected our overall results from individual studies using sensitivity analyses. For each pooled result, we conducted the following sensitivity analyses:

1. We compared each pooled result to the individual studies that contributed to the overall pooled result to determine if any studies were more influential than others and discussed this in the respective summary results narrative section. We re-ran analyses without significantly influential studies when there were two or more studies left to contribute to an overall pooled result. This was conducted in cases of significant heterogeneity or poor study quality/treatment integrity.

2. We tracked any experimental study that examined an eligible non-pharmacological treatment for infant acute pain that was excluded based on our inclusion criteria. There were 20 studies in this category (see Results). Wherever possible these studies were qualitatively described to further contextualize the overall pooled result. Results of these comparisons were included in the 'Summary of treatment effects' section at the end of each analysis but did not impact the interpretation of the quantitative analyses.

3. For every treatment analysis, we examined the studies that contributed to the overall pooled result in greater detail when they did not agree with the overall pooled result. We examined studies for methodological differences to help offer potential reasons why there was disagreement between the studies. Results of these comparisons were included in the 'Summary of treatment effects' section at the end of each analysis.

4. Despite our attempt to conduct analyses utilizing homogeneous groups of pain management interventions, age groups, and pain responses, we found substantial heterogeneity among many overall treatment effects. We still presented pooled results and we added attempts to explain heterogeneity among treatment effects to the 'Summary of treatment effects' section at the end of each treatment analysis.

This comprehensive review still provides a significant contribution to the infant pain management literature by providing treatment recommendations specifically tailored to different infant age groups and types of pain responses. Moreover, it elucidates gaps in the current literature and provides direction for future researchers in the field of non-pharmacological treatments for infant acute pain. We conducted statistical analyses using Review Manager 5.1 software (RevMan 2011).

RESULTS

Description of studies

See: Characteristics of included studies; Characteristics of excluded studies.

See the 'Characteristics of included studies' and 'Characteristics of excluded studies' tables.

Results of the search

Of the 4772 potential articles in our database search, 51 studies were included in the statistical analysis review (16 at the abstract stage; 35 after having two authors review the full paper and reach consensus for inclusion). These 51 studies had at least one treatment arm that met all the inclusion criteria of this review. We extracted means and standard deviations from the papers, or when not provided, we calculated them using the procedures outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). In addition, we referred to 20 papers, on relevant treatments that did not meet our inclusion criteria, qualitatively either within the relevant treatment analysis summary (Aguirre 2008; Bueno 2010; Campos 1989; Cignacco 2008; Cong 2009; Diego 2009; Felt 2000; Huang 2004; Johnston 2008a; Johnston 2009; Ludington-Hoe 2005; Morelius 2009; Weissman 2009) or in a separate section at the end of the results section(or both) (Campos 1989; Goubet 2003; Goubet 2007; Grunau 2004; Ipp 2009; Rattaz 2005). This allowed our recommendations to be based on the broadest possible representation of the current literature. The following description of included studies refers only to the 51 studies included in the statistical analysis.

Included studies

In the case of studies where only some of the treatment arms were included in the review, only the participants in the treatment arms that met inclusion criteria were counted towards the descriptions below.

Study design characteristics

Of the total 3396 participants that were included in the studies, 1581 were in treatment conditions only, 1153 were in control conditions, and 662 were in a cross-over study (i.e. exposure to both the treatment and control conditions). Twenty-one of the papers included in the review used cross-over designs. For cross-over designs, the participants were only counted once towards the total number of participants. The remaining 30 papers included in the review were between-groups designs.

- **Cross-over designs** (21 studies; Axelin 2006; Axelin 2009; Bellieni 2001; Bo 2000; Catelin 2005; Comaru 2009; Corbo 2000; Corff 1995; Elserafy 2009; Fearon 1997; Ferber 2008; Herrington 2007; Hill 2005; Jain 2006; Johnston 2003; Johnston 2007; Kostandy 2008; Kozub 2001; Sizun 2002; Stevens 1999; Ward-Larson 2004).
- **Between-groups designs** (30 studies; Akcan 2009; Allen 1996; Bauchner 1996; Bellieni 2002; Blass 1999; Bustos 2008; Campos 1994; Carbajal 1999; Carbajal 2003; Castral 2008; Chermont 2009; Cohen 2002; Cohen 2006; Cramer-Berness 2005; Cramer-Berness 2005b; Curtis 2007; de Sousa 2008; Gormally 2001; Gray 2000; Greenberg 2002; Hillgrove Stuart 2008; Im 2008; Ipp 2004; Johnston 1997; Kashaninia 2008;

Liaw 2010; Liu 2010; Morrow 2010; Whipple 2004; Yilmaz 2010).

Study nationality characteristics

The 51 included studies were conducted by authors in 14 different counties (United States, Canada, France, Brazil, Italy, Australia, China, Finland, Iran, Israel, Saudi Arabia, South Korea, Taiwan, and Turkey).

Description of study treatment arms

In total, the 51 included studies contributed 81 separate treatment arms (each of which was compared to a control group or control phase). We analyzed pain reactivity and pain-related regulation separately, therefore one treatment arm could be analyzed in two separate analyses if data were provided for pain reactivity (right after the pain stimulus) and pain-related regulation (occurring at least 30 seconds after the pain stimulus) separately.

- **One treatment arm:** 40 studies provided one treatment arm. Of these 40 studies, 11 studies (Axelin 2006; Catelin 2005; Chermont 2009; Corbo 2000; Ferber 2008; Gormally 2001; Johnston 2003; Kostandy 2008; Liaw 2010; Liu 2010; Yilmaz 2010) provided data for both pain reactivity and pain-related regulation and one study provided data for two age groups (Allen 1996 (neonate); Allen 1996 (older infant) for regulation. Twenty-eight studies provided data on pain reactivity or pain-related regulation (Akcan 2009; Axelin 2009; Bellieni 2001; Blass 1999; Bo 2000; Bustos 2008; Carbajal 2003; Castral 2008; Cohen 2002; Cohen 2006; Comaru 2009; Corff 1995; Curtis 2007; de Sousa 2008; Fearon 1997; Gray 2000; Greenberg 2002; Herrington 2007; Hill 2005; Ipp 2004; Jain 2006; Johnston 1997; Johnston 2007; Kashaninia 2008; Kozub 2001; Morrow 2010; Sizun 2002; Ward-Larson 2004). Thus, 53 comparisons were run in total.
- **Two treatment arms:** 10 studies provided two treatment arms. Nine of these studies contributed to 18 comparisons (Bauchner 1996; Bellieni 2002; Campos 1994; Carbajal 1999; Cramer-Berness 2005; Cramer-Berness 2005b; Im 2008; Stevens 1999; Whipple 2004). One of these 10 studies provided two arms for both pain reactivity and pain-related regulation comparisons (Hillgrove Stuart 2008; four treatment arms total). This resulted in a total of 22 comparisons with a control group/phase.
- **Three treatment arms:** one study with comparisons of both pain reactivity and pain-related regulation included three treatment arms (Elserafy 2009), resulting in six comparisons in total.

In analyses where more than one treatment arm was analyzed, we split the control group *n* to avoid the bias associated with 'double-counting'.

Description of pain responses (pain reactivity versus pain-related regulation)

A breakdown of which studies contributed to overall findings for pain reactivity (right after the pain stimulus) and pain-related regulation (occurring at least 30 seconds after the pain stimulus):

- Twenty studies reported pain reactivity outcomes, which was zero and 30 seconds after the painful stimulus (Axelin 2009; Bellieni 2001; Bellieni 2002; Bustos 2008; Carbajal 1999; Carbajal 2003; Cohen 2002; Comaru 2009; Cramer-Berness 2005; Cramer-Berness 2005b; de Sousa 2008; Hill 2005; Ipp 2004; Johnston 1997; Johnston 2007; Kashaninia 2008; Kozub 2001; Sizun 2002; Stevens 1999; Ward-Larson 2004).
- Sixteen studies reported pain-related regulation outcomes, which was the pain measurement closest to 30 seconds after the last painful stimulus (Akcan 2009; Allen 1996; Bauchner 1996; Blass 1999; Bo 2000; Campos 1994; Cohen 2006; Corff 1995; Curtis 2007; Fearon 1997; Gray 2000; Greenberg 2002; Herrington 2007; Im 2008; Jain 2006; Whipple 2004).
- Fifteen studies reported pain outcomes for both pain reactivity and pain-related regulation (Axelin 2006; Catelin 2005; Castral 2008; Chermont 2009; Corbo 2000; Elserafy 2009; Ferber 2008; Gormally 2001; Hillgrove Stuart 2008; Johnston 2003; Kostandy 2008; Liaw 2010; Liu 2010; Morrow 2010; Yilmaz 2010).

Participants

Age groupings

Of all the included studies, there were 25 studies that had preterm infants as participants (i.e. born at 36 weeks gestation or less). Sixteen papers had participants in the neonate age category (i.e. infants born at 37 weeks until one month of age). Eleven papers included older infants (i.e. over one month to 36 months of age) as participants. One paper (Allen 1996) was a cross-sectional study that included a group of neonates and older infants who underwent the same treatment, but the means and standard deviations were reported separately for each age group. For this review, the participants from this study were included separately in the neonate and older infant categories, respectively.

- **Preterm-born** (25 studies; Akcan 2009; Axelin 2006; Axelin 2009; Bellieni 2001; Castral 2008; Catelin 2005; Comaru 2009; Corbo 2000; Corff 1995; de Sousa 2008; Elserafy 2009; Fearon 1997; Ferber 2008; Herrington 2007; Hill 2005; Jain 2006; Johnston 1997; Johnston 2003; Johnston 2007; Kostandy 2008; Liaw 2010; Sizun 2002; Stevens 1999; Ward-Larson 2004; Whipple 2004).
- **Neonate** (16 studies; Allen 1996; Bo 2000; Bellieni 2002; Blass 1999; Campos 1994; Carbajal 1999; Carbajal 2003; Chermont 2009; Gormally 2001; Gray 2000; Greenberg 2002; Im 2008; Kashaninia 2008; Liu 2010; Morrow 2010; Yilmaz 2010).

- **Older infant** (11 studies; Allen 1996; Bauchner 1996; Bustos 2008; Cohen 2002; Cohen 2006; Cramer-Berness 2005; Cramer-Berness 2005b; Curtis 2007; Hillgrove Stuart 2008; Ipp 2004; Kozub 2001).

It should be noted that one study (Catelin 2005) had 66% preterm participants and 33% full-term participants. We categorized this study as a preterm study.

Diagnostic status

The diagnostic status of the infants in the 51 included studies:

- **Healthy infants born full-term** (24 studies; Allen 1996; Bauchner 1996; Bellieni 2002; Blass 1999; Bustos 2008; Campos 1994; Carbajal 1999; Carbajal 2003; Chermont 2009; Cohen 2002; Cohen 2006; Cramer-Berness 2005; Cramer-Berness 2005b; Curtis 2007; Gormally 2001; Gray 2000; Greenberg 2002; Hillgrove Stuart 2008; Im 2008; Ipp 2004; Kashaninia 2008; Liu 2010; Morrow 2010; Yilmaz 2010).
- **Hospitalized infants born preterm** (25 studies; Axelin 2006; Akcan 2009; Axelin 2009; Bellieni 2001; Castral 2008; Catelin 2005; Comaru 2009; Corbo 2000; Corff 1995; de Sousa 2008; Elserafy 2009; Fearon 1997; Ferber 2008; Herrington 2007; Hill 2005; Jain 2006; Johnston 1997; Johnston 2003; Johnston 2007; Kostandy 2008; Liaw 2010; Sizun 2002; Stevens 1999; Ward-Larson 2004; Whipple 2004).
- **Infants born full-term, being monitored or treated for health complications** (two studies; Bo 2000; Kozub 2001).

Types of painful procedures

For a study to be included in the search, the procedure had to be described by the author as painful or causing pain (e.g. diaper change or weighing procedure for preterms). The following painful procedures were used in the included studies:

- **Heel-stick** (29 studies; Akcan 2009; Axelin 2009; Bellieni 2001; Bellieni 2002; Blass 1999; Bo 2000; Campos 1994; Castral 2008; Corbo 2000; Corff 1995; de Sousa 2008; Fearon 1997; Ferber 2008; Gormally 2001; Gray 2000; Greenberg 2002; Herrington 2007; Im 2008; Jain 2006; Johnston 1997; Johnston 2003; Johnston 2007; Kostandy 2008; Kozub 2001; Liaw 2010; Morrow 2010; Stevens 1999; Whipple 2004;Yilmaz 2010).
- **Vaccine/vitamin needle** (10 studies; Allen 1996; Bustos 2008; Chermont 2009; Cohen 2002; Cohen 2006; Cramer-Berness 2005; Cramer-Berness 2005b; Hillgrove Stuart 2008; Ipp 2004; Kashaninia 2008).
- **Venipuncture** (six studies; Bauchner 1996; Carbajal 1999; Carbajal 2003; Curtis 2007; Elserafy 2009; Liu 2010).
- **Diaper change** (two studies; Comaru 2009; Sizun 2002).
- **Endotracheal suctioning technique** (two studies; Axelin 2006; Ward-Larson 2004).
- **Weighing procedure** (two studies; Catelin 2005, Hill 2005).

Types of treatment

In total, we analyzed 13 different types of treatment for this review. We analyzed the efficacy of the 13 different types of treatment separately for preterms, neonates, and older infants. Moreover, within each treatment X age analysis, data were separated into pain types (pain reactivity or pain-related regulation). None of the 81 treatment arms from the 51 included studies were included twice within any treatment X age X pain type analyses. Eleven studies contributed more than one treatment arm. They are distinguished by the suffix "-1", "-2" or "-3" added to the in-text citation reference. Noteworthy comments regarding the general grouping of the 35 treatment comparisons follow.

We grouped studies that used pacifiers and other methods of non-nutritive sucking together under 'Non-nutritive sucking-related'. Due to a number of studies that included water as a treatment arm (thus enabling a comparison with an actual 'no treatment' arm), this was included as a separate treatment, despite the fact that some studies considered the administration of water (without a sucking tool such as a pacifier) as a control group. We grouped 'swaddling/facilitated tucking' together due to sharing one similar mechanism of comfort (containment), despite having differences in other pain-reducing qualities beyond that of containment. Finally, 'structured parent involvement' were interventions that were grouped together if parents were given some type of instructive suggestion (whether written or oral) on how to reduce pain for their infant with no other tools given.

1. Kangaroo care (14 treatment arms)
   - i) Preterm
     - a) *Pain reactivity*: Castral 2008; de Sousa 2008; Ferber 2008; Johnston 2003; Kostandy 2008
     - b) *Pain-related regulation*: Akcan 2009; Castral 2008; Ferber 2008; Johnston 2003; Kostandy 2008
   - ii) Neonate
     - a) *Pain reactivity*: Chermont 2009; Kashaninia 2008
     - b) *Pain-related regulation*: Chermont 2009; Gray 2000
   - iii) Older infant: no studies found on kangaroo care for older infants
2. Swaddling or tucking (10 treatment arms)
   - i) Preterm
     - a) *Pain reactivity*: Axelin 2006; Axelin 2009; Comaru 2009-1; Hill 2005; Stevens 1999; Ward-Larson 2004
     - b) *Pain-related regulation*: Axelin 2006; Corff 1995; Fearon 1997
   - ii) Neonate
     - a) Pain reactivity: Morrow 2010
   - iii) Older infant: no studies found for swaddling/tucking for older infants
3. Non-nutritive sucking-related (24 treatment arms)
   - i) Preterm
     - a) *Pain reactivity*: Corbo 2000; Bellieni 2001; Elserafy 2009-1; Elserafy 2009-2; Liaw 2010; Stevens 1999

b) *Pain-related regulation*: Corbo 2000; Elserafy 2009-1; Elserafy 2009-2; Liaw 2010; Whipple 2004-1; Whipple 2004-2

ii) Neonate

a) *Pain reactivity*: Bellieni 2002; Carbajal 1999; Liu 2010; Yilmaz 2010

b) *Pain-related regulation*: Blass 1999; Bo 2000; Campos 1994; Greenberg 2002; Im 2008; Liu 2010; Yilmaz 2010

iii) Older infant

a) *Pain reactivity*: no studies found

b) *Pain-related regulation*: Curtis 2007

4. Swallowing water (five treatment arms)

i) Preterm

a) *Pain reactivity*: Elserafy 2009-3

b) *Pain-related regulation*: Elserafy 2009-3

ii) Neonate

a) *Pain reactivity*: Carbajal 1999

b) *Pain-related regulation*: Allen 1996

iii) Older infant

a) *Pain reactivity*: no studies found

b) *Pain-related regulation*: Allen 1996

5. Rocking, holding or both (five treatment arms)

i) Preterm: no studies found for rocking/holding for preterms

ii) Neonate

a) *Pain reactivity*:Carbajal 2003; Gormally 2001

b) *Pain-related regulation*: Campos 1994; Gormally 2001

iii) Older infant

a) *Pain reactivity*: Ipp 2004

b) *Pain-related regulation*: no studies found for rocking/holding for older infants

6. Simulated rocking and water(one treatment arm)

i) Preterm

a) *Pain reactivity*: Johnston 1997

b) *Pain-related regulation*: no studies found

ii) Neonate: no studies found for simulated rocking for neonates

iii) Older infant: no studies found for simulated rocking for older infants

7. Touch or massage (five treatment arms)

i) Preterm

a) *Pain reactivity*: no studies found

b) *Pain-related regulation*: Herrington 2007; Jain 2006

ii) Neonate

a) *Pain reactivity*: Bellieni 2002

b) *Pain-related regulation*: Im 2008

iii) Older infant

a) *Pain reactivity*: Kozub 2001

b) *Pain-related regulation*: no studies found

8. Environmental modification (three treatment arms)

i) Preterm

a) *Pain reactivity*: Catelin 2005; Sizun 2002

b) *Pain-related regulation*: Catelin 2005

ii) Neonate: no studies found for environmental modification for neonates

iii) Older infant: no studies found for environmental modification for older infants

9. Toy distraction (six treatment arms)

i) Preterm: no studies found for toy distraction for preterm infants

ii) Neonate: no studies found for toy distraction for neonates

iii) Older infant

a) *Pain reactivity*: Cramer-Berness 2005; Cramer-Berness 2005b; Hillgrove Stuart 2008-1; Hillgrove Stuart 2008-2

b) *Pain-related regulation*: Hillgrove Stuart 2008-1; Hillgrove Stuart 2008-2

10. Video distraction (two treatment arms)

i) Preterm: no studies found for video distraction for preterm infants

ii) Neonate: no studies found for video distraction for neonates

iii) Older infant

a) *Pain reactivity*: Cohen 2002

b) *Pain-related regulation*: Cohen 2006

11. Structured parent involvement (four treatment arms)

i) Preterm: no studies found for structured parent involvement for preterm infants

ii) Neonate: no studies found for structured parent involvement for neonates

iii) Older infant

a) *Pain reactivity*: Bustos 2008; Cramer-Berness 2005; Cramer-Berness 2005b

b) *Pain-related regulation*: Bauchner 1996

12. Maternal voice (one treatment arm)

i) Preterm

a) *Pain reactivity*: Johnston 2007

b) *Pain-related regulation*: no studies found

ii) Neonate: no studies found for maternal voice for neonates

iii) Older infant: no studies found for maternal voice for older infants

13. Parent presence (one treatment arm)

i) Preterm: no studies found for parent presence for preterm infants

ii) Neonate: no studies found for parent presence for neonates

iii) Older infant

a) *Pain reactivity*: no studies found

b) *Pain-related regulation*: Bauchner 1996

Treatment locations

The treatment interventions described in the studies occurred in the following locations:

- **Inpatient hospital** (44 studies; Akcan 2009; Allen 1996; Axelin 2006; Axelin 2009; Bauchner 1996; Bellieni 2001; Bellieni 2002; Blass 1999; Bo 2000; Campos 1994; Castral 2008; Carbajal 1999; Carbajal 2003; Catelin 2005; Chermont 2009; Comaru 2009; Corbo 2000; Corff 1995; Curtis 2007; de Sousa 2008; Elserafy 2009; Fearon 1997; Ferber 2008; Gormally 2001; Gray 2000; Greenberg 2002; Herrington 2007; Hill 2005; Im 2008; Jain 2006; Johnston 1997; Johnston 2003; Johnston 2007; Kashaninia 2008; Kostandy 2008; Kozub 2001; Liaw 2010; Liu 2010; Morrow 2010; Sizun 2002; Stevens 1999; Ward-Larson 2004; Whipple 2004; Yilmaz 2010).
- **Outpatient medical clinic** (seven studies; Bustos 2008; Cohen 2002; Cohen 2006; Cramer-Berness 2005; Cramer-Berness 2005b; Hillgrove Stuart 2008; Ipp 2004).

Excluded studies

We retrieved 4772 abstracts using the search strategies described in the protocol. From these titles/abstracts, 133 articles were reviewed and 16 were automatically included and 4639 were automatically excluded (RPR, AG, LU, KT). We used the remaining 117 abstracts to find papers that required further examination to make a decision (RPR, NR, KT, LU, RH, LDO, SAK, JHS). Two review authors reviewed these papers to determine whether they met the inclusion criteria for the current review (RPR, KT). Of these 117 studies, we excluded 62 studies. These were the primary reasons for exclusion for the 62 studies that needed to be reviewed:

- **Inappropriate age or infant age group could not be separated from older child group** (36 studies; Abedin 2008; Al-Bekaa 2003; Aslanabadi 2008; Carlson 2000; D'Agostino 2008; Dahlquist 2002; Drago 2009; Favara-Scacco 2001; Gedaly-Duff 1992; Gold 2006; Gonzalez 1989; Gonzalez 1993; Hatem 2006; Heden 2009; Hoffman 2006; Ida 2008; Jackson 2008; Jo 2007; Kivijärvi 2008; Koivusalo 2009; Li 2007; MacLaren 2005; Manne 1990; Marchisotti 2007; Marec-Berard 2009; McCarthy 2010; Michel 2008; Phipps 2005; Reichel 2007; Salmon 2006; Shapiro 2007; Slifer 2009; Sparks 2007; Sundararajan 2007; Tanabe 2002; Zeltzer 1991).
- **Inappropriate intervention** (14 studies; Axelin 2010; Bellieni 2003; Bellieni 2007; Boots 2010; Dilen 2010; El-Naggar 2010; Hanson 2010; He 2010; Holsti 2005; Ipp 2007; Mucignat 2004; Ozdogan 2010; Slater 2010; Wisdorf-Houtkooper 1997).
- **No pain or pain not measured on day of procedure** (nine studies; Cologna 1999; Duncan 2004; Harrison 2000; Hsu 1995; Huang 1999; Johnston 2007b; Leclair 2007; Marin Gabriel 2010; Vignochi 2010).
- **Student work later published and included in review** (one study; Greenberg 1997).
- **Randomized controlled trial in progress** (one study; Campbell-Yeo 2009).
- **Translation issues and could not contact author** (one study; Park 2006).

We retrieved one article in another language (Korean) and had a native Korean speaker unsuccessfully attempt to extract the needed means and standard deviations (Park 2006). We were unable to retrieve the necessary information and the author could not be contacted, resulting in the exclusion of this article from the review. In addition to the above reasons for excluded articles, 20 studies were in a gray area as they did not meet the full inclusion criteria to be included in the statistical analyses of the review, but it was felt they were relevant to the purpose of the review. These papers were referred to either in the 'Summary of treatment effects' narratives following the relevant quantitative analysis of the included studies OR they were included in a separate section at the end of the results ('Other potentially effective non-pharmacological interventions') when there was no relevant quantitative intervention analysis for comparison. In this way, qualitative sensitivity analyses could be conducted where the summary outcome of the quantitative analyses could be compared to the results of individual studies that did not meet the inclusion criteria. These studies did not meet complete inclusion criteria for the following reasons:

- **Control group was an active control/no control group** (14 studies; Aguirre 2008; Bueno 2010; Campos 1989; Cignacco 2008; Diego 2009; Goubet 2003; Goubet 2007; Grunau 2004; Huang 2004; Ipp 2009; Johnston 2008a; Johnston 2009; Ludington-Hoe 2005; Rattaz 2005).
- **No behavioral pain outcome reported or not analyzable** (two studies; Cong 2009; Okan 2010).
- **Was not a RCT** (four studies; Felt 2000; Morelius 2009; Vivancos 2010; Weissman 2009).

Risk of bias in included studies

Six raters, in three groups of two, participated in study quality rating (RPR, NR, RH, JHS, SAK, LDO). Each team was assigned a subset of the 51 articles included in the quantitative analysis. Two review authors independently rated each article. Every individual score on the data extraction form (article details, study quality, and treatment integrity) were compared by the two raters within a team. Every score that differed between the two members of the team was examined. The data extraction form required raters to put the page and paragraph number (from the original article) that justified their rating. Thus, when scores differed between team members, the original article was consulted and a consensus-based decision was made.

Study quality

While any cut-off is of an arbitrary nature, for risk of bias analysis purposes, it was determined that studies would be classified

into one of two categories. These two categories were based on the following reasoning: a) a natural 'break' in the distribution of total scores occurred at 13; and b) a post-hoc examination of mean scores on the key sections of the risk of bias scale (i.e. "4. Steps to Reduce Bias"; "5. Use of Valid and Reliable Outcome Measures"; "6. Adequate Statistical Analyses") indicated that there were consistent differences between the lower/unknown and satisfactory/marginal studies in the expected directions (i.e. the means of satisfactory/marginal quality studies were always higher than the lower/unknown quality studies). It should be noted, however, that low scores do not necessarily reflect poor methodological quality but rather often reflected suboptimal reporting practices, especially in publications published before/around the CONSORT guidelines became widely disseminated (circa 1996).

Thus, we considered studies which had scores of 14 or higher (out of 23) of satisfactory/marginal quality and we classified studies with scores of 13 or less as lower/unknown quality studies. We considered 90% of studies of satisfactory/marginal quality (46 studies) and the remaining studies were in the lower/unknown category.

Importantly, we noted the studies that were rated of lower/unknown quality for the results and authors conclusion sections. Five of the 35 comparisons included treatment arms from studies that were of lower/unknown quality. We re-ran these four analyses without the studies of questionable quality to determine how our results would have differed. The impact of removing these studies has been included in the respective summaries in the 'Effects of interventions' section. Where applicable, we qualified treatment recommendations with a caveat regarding the impact of lower/unknown quality studies.

- **Satisfactory/marginal quality** (46 studies; Akcan 2009; Axelin 2006; Axelin 2009; Bauchner 1996; Bellieni 2001; Bellieni 2002; Blass 1999; Bo 2000; Bustos 2008; Castral 2008; Carbajal 1999; Carbajal 2003; Catelin 2005; Chermont 2009; Cohen 2006; Comaru 2009; Corbo 2000; Corff 1995; Cramer-Berness 2005; Cramer-Berness 2005b; Curtis 2007; de Sousa 2008; Elserafy 2009; Fearon 1997; Ferber 2008; Gormally 2001; Gray 2000; Herrington 2007; Hill 2005; Hillgrove Stuart 2008; Im 2008; Ipp 2004; Jain 2006; Johnston 1997; Johnston 2003; Johnston 2007; Kashaninia 2008; Kostandy 2008; Kozub 2001; Liaw 2010; Liu 2010; Morrow 2010; Sizun 2002; Stevens 1999; Ward-Larson 2004; Yilmaz 2010).
- **Lower/unknown quality** (five studies; Allen 1996; Campos 1994; Cohen 2002; Greenberg 2002; Whipple 2004).

To further contextualize the overall quality of the studies included in this review, summary descriptions have been provided below for each of the seven sections that comprised our study quality tool. Following these descriptions is a summary of our findings regarding the treatment integrity of the studies.

1. Inclusion/exclusion criteria clearly specified: overall, this was done extremely well with 96% of studies adequately describing sample criteria and 88% of studies providing some type of quantitative or qualitative information to inform readers that the sample recruitment criteria was met.

2. Attrition reported according to CONSORT guidelines: while this question automatically biases quality ratings against studies published around 1996 (pre-CONSORT), only four studies were published before this time period. While 47% provided sufficient documentation of attrition, they did not explicitly mention CONSORT, while 29% did both. Thirty-eight percent of studies did not provide adequate information regarding attrition.

3. Good sample description: overall, this was another area of strength for the studies included in this review with 88% providing a satisfactory description of the sample and 88% providing some indication that the samples in the treatment groups were equivalent.

4. Steps to minimize bias: this section dealt with reporting of randomization, allocation sequence, and measurement bias. While all studies purported a randomized assignment of participants only 37% provided data that allowed us to surmise a convincing method for generating random participant assignment and the use of person not involved in recruitment or running the study. Forty-five percent described good randomization practices but not the use of an independent person. The majority of studies reported an allocation method that ensured the assignment of participants was blinded to the investigators (63%). Finally, approximately half the studies (47%) explicitly reported that data collectors/coders were blinded to the patient group. It should be noted that this final point on measurement bias is a very challenging standard for studies that compare a non-pharmacological interventions (usually entailing overt behaviors such as kangaroo care or exposing child to a toy) to a no intervention control group or phase.

5. Use of valid and reliable outcome measures: overall, all the studies utilized measures that would be considered justified as a measure of pain-related distress in infancy. However, 26% of studies did not provide information where a clear justification of the measures, in the context of the treatment aims, could be discerned and 22% of studies used at least one measure that was not valid or known to be valid in the context of the study (however, again, the measure of pain-related distress was considered valid for all studies). Finally, 98% of studies reported data to demonstrate that the majority of outcome measures were reliably measured and sensitive to change.

6. Adequate statistical analyses: this section involved five different standards to get full points: using a power analysis, having a sufficient sample size (based on power analysis if null findings), adequate analyses for a treatment arm compared to a control group, adequate reporting of summary statistics, and an intention-to-treat analysis. On the whole, the majority of studies demonstrated evidence of adequate sample size (75%), adequate analysis plan (98%), and reporting of summary statistics (80%). However, only 63% of studies reported conducting a power

analysis and only 47% of studies reported intention-to-treat analysis. This low 'intention-to-treat' percentage should be qualified with the fact that in the majority of these studies, most, if not all randomized participants were analyzed. Our high participant randomization to participant analysis ratio is partly due to the fact that we analyzed behavioral data only. A major reason for excluding physiological data was the significant proportions of participants that were excluded due to missing data.

7. Well-matched control group: given the focus of the review on randomized non-pharmacological intervention studies we were advised that 'active control groups' needed to be excluded from analysis as our primary questions related to the value of interventions compared to no interventions. Thus, having a control group that was well matched in terms of structural features of the treatment was not considered a necessary standard. When it was deemed possible, to get full points, review authors needed to be able to discern that there were reasonable similarities between treatment and no treatment phases (e.g. infants in either group received same painful stimulus in same environment by same professionals). All our cross-over studies met this standard (21 studies) and 21/30 of our parallel studies did not receive full credit. The majority of times a lack of reporting issue (i.e. no explicit mention of the equivalence of intervention to no intervention) was suspected rather than the study itself not having a good well-matched control group.

Treatment integrity

Overall, the integrity of the 51 studies were analyzed using a qualitative scale. Five areas were considered when making the summary judgment (Higgins 2011; treatment adherence, treatment exposure, quality of delivery, participant responsiveness and program differentiation). Of the 51 studies, we classified 50 'satisfactory/good' and one was 'poor'. The study rated as 'poor' lost points across all five subscales used to make the overall judgment. The quantitative analysis was run with and without this article to determine if it impacted the overall direction of the results.

- **Satisfactory/good** (50 studies; Akcan 2009; Allen 1996; Axelin 2006; Axelin 2009; Bauchner 1996; Bellieni 2001; Bellieni 2002; Blass 1999; Bo 2000; Bustos 2008; Campos 1994; Carbajal 1999; Carbajal 2003; Castral 2008; Catelin 2005; Chermont 2009; Cohen 2002; Cohen 2006; Comaru 2009; Corbo 2000; Cramer-Berness 2005; Cramer-Berness 2005b; Curtis 2007; de Sousa 2008; Elserafy 2009; Fearon 1997; Ferber 2008; Gormally 2001; Gray 2000; Greenberg 2002; Herrington 2007; Hill 2005; Hillgrove Stuart 2008; Im 2008; Ipp 2004; Jain 2006; Johnston 1997; Johnston 2003; Johnston 2007; Kashaninia 2008; Kostandy 2008; Kozub 2001; Liaw 2010; Liu 2010; Morrow 2010; Sizun 2002; Stevens 1999; Ward-Larson 2004; Whipple 2004; Yilmaz 2010).
- **Poor** (one study; Corff 1995).

Effects of interventions

To analyze the efficacy of each intervention (see 'Types of treatment' above), we first separated the studies within an intervention category by the age of the sample (preterm: [infants born less than 37 weeks gestation; hospitalized after birth], or neonate [full-term infants between birth and one month of age], or older infant [full-term infants; aged from one month to 36 months of age]).

We further subdivided outcomes by the timing of the pain response that was being measured, either pain reactivity (within the first 30 seconds following the painful stimuli) or pain-related regulation (the next closest time point measured 30 seconds after the painful stimuli). For example, for studies that examined sucking-related interventions, first we divided into three age groups (preterm, neonate or older infant), then we divided by timing of response measurement (pain reactivity or pain-related regulation). Thus, treatment recommendations were age-specific and pain response-specific (e.g. sucking-related treatment recommendation for preterm infants' pain reactivity), as opposed to a blanket recommendation for sucking-related interventions.

Due to the emerging verbal skills of infants and recommended assessment procedures for infant pain (Franck 2000; Stevens 2007), we made the decision to only analyze objectively measured behavioral responses to pain (see 'Types of outcome measures'). When studies had more than one behavioral response to pain, we used the most valid measure available. Specifically, we used pain facial expression over any other behavior. While initially we had planned to analyze physiological data, upon review of the literature, physiological measures were not included due to considerable methodological heterogeneity (including significant missing data that were present for behavioral measures; in these situations it was rare that intention-to-treat analyses were done). However, if well-established, multi-dimensional pain measures were utilized, with behavioral and physiological indicators, and it was a reliable and valid measure, we used the total score (for example, the Premature Infant Pain Profile (PIPP; Stevens 1996). Three studies utilized measures of distress that directly followed a painful procedure, and thus were included as a measure of pain (Corbo 2000; Ferber 2008; Whipple 2004).

We standardized all the scores using the standardized mean difference (SMD) method recommended by The Cochrane Collaboration. Since all the outcomes were behaviorally-based and standardized but used different types of measurements, we examined them together as either the outcome variable called 'Pain-related distress reactivity' (pain reactivity) or 'Pain-related distress regulation' (pain-related regulation). As such, our outcome measure was treated the same for each intervention and it encapsulated all the above measures of pain-related distress.

Standardized mean differences (SMDs) using a random-effects model are displayed in the results below with the confidence intervals included in brackets. Interventions for specific ages and time periods are considered effective when the SMD and the two anchors of the confidence interval fell in the negative range. All

participants included in the analyses were randomized and met the inclusion criteria outlined earlier. Studies that examined a relevant non-pharmacological treatment but did not meet the inclusion criteria were qualitatively compared to the actual quantitative analyses in the individual 'Summary of treatment effects' section. In terms of interpretation of the effect sizes, 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 a large effect (Cohen 1988). For this review we used the inverse variance method recommended by The Cochrane Collaboration. Using this method, the weight given to each study is the inverse of the variance of the effect estimate. This means that larger studies with smaller standard errors are given more weight than smaller studies with larger standard errors. Thus, since the overall findings are heavily impacted by heterogeneity (i.e. large variance in treatment or control groups), the overall results presented may differ from the results of individual studies. Also, as aforementioned, based on our study quality analyses, studies of lower/unknown quality were examined for their impact on the findings for individual treatments.

The results below examine the effects of 13 different treatment interventions. In total, there are 35 separate comparisons (13 preterm infant comparisons (seven for pain reactivity, six for immediate regulation), 11 neonate comparisons (six for pain reactivity, five for immediate regulation), and 11 comparisons for older infants (five for pain reactivity, six for immediate regulation)). The individual 'Summary of results' tables indicate the number of treatment arms included in a comparison and the total number of participants included in the comparison (including participants in the control group). Please refer to 'Table 2' for an overall summary of the entire review. See Additional tables 3-15

1. Kangaroo care

(Please see 'Summary of results Table 3' for SMDs and CIs for the effect of kangaroo care on pain-related distress reactivity and immediate regulation).

i) Preterm infants

a. Pain reactivity

For preterm infants' pain reactivity, five treatment arms contributed to the finding that kangaroo care was significantly better than a no-treatment control group, with a SMD of -1.12 (95% CI -2.04 to -0.21). See Analysis 1.1 for the full analysis. However, the heterogeneity statistic (I^2 = 89%) suggested substantial heterogeneity among the treatment effects of the analyzed studies. Two studies contributed to the I^2 statistic due to larger than average SMDs (de Sousa 2008; Ferber 2008). With these two studies removed, the SMD became -0.38 (95% CI -0.65 to -0.12; I^2 = 0%).

b. Immediate regulation

Results from five treatment arms found that kangaroo care was significantly superior to control, with a SMD of -0.77 (95% CI -1.50 to -0.03), for the immediate regulation of pain-related distress in preterm infants. See Analysis 2.1 for the full analysis. However, the heterogeneity statistic (I^2 = 82%) suggested substantial heterogeneity among the treatment effects of the analyzed studies. One study contributed to the I^2 statistic due to a larger than average SMD (Ferber 2008). With this study removed, the SMD became -0.45 (95% CI -0.69 to -0.20; I^2 = 0%), remaining significant.

ii) Neonates

a. Pain reactivity

Although the SMD for the effect of kangaroo care on pain reactivity in neonates fell in the negative range (-0.89), its CI included 0 (95% CI -2.89 to 1.10), indicating that kangaroo care may not significantly differ from a no-treatment control group. Only two studies were included in this sub-analyses and the heterogeneity quotient was substantial (I^2 = 98%), indicating that the comparability of the two treatment effects may be questionable. See Analysis 3.1 for the full analysis.

b. Immediate regulation

The SMD for the effect of kangaroo care fell in the negative range (-0.66), but the CI included 0 (95% CI -1.73 to 0.42). Kangaroo care did not differ from a no-treatment control for immediate regulation of pain-related distress in neonates (see Analysis 4.1 for the full analysis). It should be noted that the heterogeneity quotient suggested substantial heterogeneity, thus, comparability of the two studies' results may be questionable (I^2 = 82%).

Summary of treatment effect: kangaroo care

Preterm infants. The results from this review found that kangaroo care is efficacious in reducing pain reactivity and improving pain-related regulation for preterm infants. All studies were of satisfactory/marginal quality. While there was substantial variability among the treatment effects, the overall SMD is evidence to recommend kangaroo care as an intervention for reducing pain reactivity and improving regulation of pain-related distress in preterm infants. When the analysis were re-run without the sources of heterogeneity, results remained consistent. Of peripheral interest, four studies that were not included in the statistical analyses due to not meeting our methodological criteria (Cong 2009; Johnston 2008a; Johnston 2009; Ludington-Hoe 2005) also lend support that kangaroo care may be effective in improving pain reactivity and regulation of pain-related distress in preterm infants.

Neonates. For neonates, the results suggest that kangaroo care is not efficacious as an intervention for pain reactivity nor pain-

related regulation. All studies were of satisfactory/marginal quality. However, based on an examination of the methodologies in the studies and the significant treatment effect for preterms, future research should explore whether using a longer exposure time in kangaroo care prior to the painful procedure (i.e. 10 minutes or longer for neonates), could lead to a significant treatment effect.

2. Swaddling/tucking-related

(Please see 'Summary of results Table 4' for SMDs and CIs for the effect of swaddling/tucking on pain-related distress reactivity and immediate regulation).

i) Preterm infants

a. Pain reactivity

Six treatment arms investigated the effect of swaddling or tucking on pain-related distress pain reactivity in preterm infants. One study reported two groups separately who both received tucking and nesting (Comaru 2009) with separate control groups. These two groups were thus entered into the analysis as separate studies since they used different participants. Swaddling/tucking was found to be effective in reducing pain-related distress pain reactivity in preterm infants (SMD -0.97; 95% CI -1.63 to -0.31). See Analysis 5.1 for the full analysis. However, it should be noted that the heterogeneity coefficient was substantial (I^2 = 88%). Two studies contributed to the I^2 statistic due to a greater than average absolute SMD (Stevens 1999; Ward-Larson 2004; I^2 = 0%). With these studies removed, the SMD became -0.90 (95% CI -1.22 to -0.59; I^2 = 0%), remaining significant.

b. Immediate regulation

Three studies investigated whether facilitated tucking was effective for the immediate regulation of pain-related distress of preterm infants. Facilitated tucking was an efficacious intervention for improving the immediate regulation of preterm infants (SMD -0.75; 95% CI -1.14 to -0.36). See Analysis 6.1 for the full analysis. There was not substantial heterogeneity among studies. Of further note, the Corff 1995 study was deemed to have poor treatment integrity and when the treatment effects were analyzed without this study, the overall effect for immediate regulation of pain-related distress stayed significant (SMD -0.61; 95% CI -1.12 to -0.11; I^2 = 0).

ii) Neonates

a) Pain reactivity

Results from one small study suggest that swaddling/tucking is effective for pain-related distress with a SMD of -1.26 (95% CI -1.92 to -0.60). See Analysis 7.1 for the full analysis.

Summary of treatment effect: swaddling/tucking-related

Preterm infants. There was sufficient evidence to support the use of swaddling/tucking as an effective intervention for reducing pain-related distress reactivity and regulation in preterm infants. Of side note, two studies (Huang 2004; Johnston 2008a) that were not included in the analysis due to use of an active control group, suggested that swaddling was as effective as containment (placing the child in a lateral position for 30 minutes prior to heel-stick) but not as effective as kangaroo care.
Neonates. Limited evidence supports the effectiveness of swaddling/tucking-related interventions for the healthy neonate.

3. Non-nutritive sucking-related

(Please see 'Summary of results Table 5' for SMDs and CIs for the effect of non-nutritive sucking on pain-related distress reactivity and immediate regulation).

i) Preterm infants

a. Pain reactivity

Five studies (6 treatment arms) investigated the effects of sucking on preterm pain-related distress pain reactivity (SMD -0.42; 95% CI -0.68 to -0.15). One study (Elserafy 2009) contributed two treatment arms to the analysis: a pacifier treatment arm and a pacifier plus water treatment arm. See Analysis 8.1 for the full analysis. High heterogeneity was determined (I^2 = 48%) due to the presence of one larger SMD (Corbo 2000). When analyses were re-run without this study the results stayed significant (SMD -0.32; 95% CI -0.5 to -0.15; I^2 = 0%).

b. Immediate regulation

Four studies investigated the effects of sucking on preterm pain-related regulation, with two studies contributing two arms each (Elserafy 2009; Whipple 2004). For the arms included in the analysis, four involved sucking on a pacifier, one involved sucking on a pacifier with water, and one involved sucking on a pacifier with music. There is sufficient evidence that sucking is efficacious for improving immediate regulation in preterm infants (SMD -0.38; 95% CI -0.59 to -0.17). Of side note, a study not included in the quantitative analyses due to use of an active control group (Campos 1989) also suggests sucking is an effective intervention, when compared to swaddling. See Analysis 9.1 for full analysis. In addition, when analyses were re-run without the study of lower quality (Whipple 2004) significance of the treatment effect was maintained (SMD -0.36; 95% CI -0.59 to -0.13; I^2 = 0%).

ii) Neonates

a. Pain reactivity

Four studies investigated the effect of sucking on pain-related distress pain reactivity for neonates (see Analysis 10.1). The SMD of -1.45 (95% CI -2.34 to -0.57; I^2 = 88%) suggests that for neonates, sucking is efficacious in reducing immediate pain reactivity. One study involved stimulating sucking with water for 30 seconds after the needle, where as the other involved placing a pacifier in the baby's mouth. When analyses were re-run without the studies contributing to the substantial heterogeneity (Yilmaz 2010), the effect size remained significant (SMD -1.88; 95% CI -2.25 to -1.50; I^2 = 0%)

b. Immediate regulation

Seven studies investigated the effect of sucking on the immediate regulation of pain-related distress of neonates (see Analysis 11.1). An overall SMD of -0.90 (95% CI -1.54 to -0.25; I^2 = 84%) suggests there is sufficient evidence that sucking is efficacious for improving immediate regulation in neonates. When analyses were re-run without studies that had notably different absolute SMDs and/or lower/unknown quality (Bo 2000; Greenberg 2002; Liu 2010), the results remained significant (SMD -0.60; 95% CI -0.91 to -0.29; I^2 = 11%).

iii) Older infants

a. Immediate regulation

Only one study with 41 participants investigated the effect of sucking on the immediate regulation of pain-related distress of older infants. This study found that sucking is an efficacious intervention for helping older infants to regulate following a painful procedure (SMD -0.89; 95% CI -1.53 to -0.25). Thus, sucking-related interventions are considered a promising intervention with confirmatory evidence needed.

Summary of treatment effect: non-nutritive sucking-related

Preterm infants. The pooled results from this review suggest that there is sufficient evidence that sucking is efficacious in reducing pain-related distress reactivity and regulation for preterm infants. An analysis of significant studies suggests that pain relief will be maximized if sucking begins at least three minutes prior to the painful stimuli. Of peripheral note, two studies that were not included in the analyses due to the use of an active control group (Bueno 2010; Weissman 2009) also suggest that sucking helps diminish pain-related distress pain reactivity in preterm infants.

Neonates. The results show sufficient evidence for sucking to moderate pain-related reactivity and immediate regulation in neonates. Four studies that were not included in the analyses, due to methodological differences with our included studies (Aguirre 2008; Bueno 2010; Campos 1989; Morelius 2009) also lend support to the efficacy of sucking to improve pain-related regulation for neonates. There is sufficient evidence that sucking is more efficacious than a no-treatment control for reducing pain-related distress pain reactivity in preterm infants.

Older infants. The results of one small satisfactory/marginal quality study suggest that sucking has limited evidence as an efficacious intervention to improve pain-related distress pain reactivity for older infants.

4. Swallowing water

(Please see 'Summary of results Table 6' for SMDs and CIs for the effect of swallowing water on pain-related distress reactivity and immediate regulation).

i) Preterm infants

a. Pain reactivity

The effect of swallowing water on pain-related distress pain reactivity in preterm infants was assessed in one small study. This study found that water was not more efficacious than a no-treatment control in reducing preterm pain-related distress pain reactivity (SMD -0.24; 95% CI -0.71 to 0.23).

b. Immediate regulation

One small study examined the effect of swallowing water on the immediate regulation of pain-related distress of preterm infants. Based on this result, water was not an efficacious intervention for reducing pain-related distress upon immediate regulation for preterm infants (SMD -0.23; 95% CI -0.70 to 0.24).

ii) Neonates

a. Pain reactivity

The effect of swallowing water on pain-related distress pain reactivity in neonates was assessed in one small study. This study found that water was not efficacious in reducing pain-related distress pain reactivity for neonates (SMD 0.10; 95% CI -0.45 to 0.66).

b. Immediate regulation

One small study examined the effect of swallowing water on the immediate regulation of pain-related distress of neonates. Based on this result, water is not an efficacious intervention for reducing pain-related distress upon immediate regulation for neonates (SMD -0.53; 95% CI -1.21 to 0.16).

iii) Older infants

a. Immediate regulation

One small study examined the effect of swallowing water on the immediate regulation of pain-related distress of older infants. Based on the results of one study, water is not more efficacious than a no-treatment control for reducing pain-related distress upon immediate regulation for older infants (SMD 0.00; 95% CI -0.72 to 0.72).

Summary of treatment effects: swallowing water

Preterm infants. There is limited evidence that water is an ineffective intervention for pain reactivity or pain-related regulation for preterm infants.
Neonates. There is limited evidence that water is an ineffective intervention for pain reactivity or pain-related regulation of neonates. The study for pain-related regulation was of low/unknown quality.
Older infants. There is limited evidence that water is an ineffective intervention for pain-related regulation for older infants. The study that found no effect was of low/unknown quality.
The above studies used 'water' as a treatment arm (in comparison to a 'no-treatment' control), while most other studies in the literature used water as the 'no-treatment' control group. Given the more common use of water in the literature and the limited evidence at every age group of its inefficacy, it is not recommended that further research explores water as a treatment for young child procedural pain.

5. Rocking/holding

(Please see 'Summary of results Table 7' for SMDs and CIs for the effect of rocking/holding on pain-related distress reactivity and immediate regulation).

i) Neonates

a. Pain reactivity

Two studies investigated the effect of holding on the pain-related distress pain reactivity of neonates following a painful procedure. The pooled effects found that rocking/holding alone is not more effective than a no-treatment control group in reducing pain-related distress pain reactivity in neonates (SMD -0.33; 95% CI -1.05 to 0.39). The heterogeneity quotient was substantial (I^2 = 73%), with one study finding no effect and the other finding that there was a significant SMD. See Analysis 18.1 for the complete analysis.

b. Immediate regulation

The effect of rocking or holding on immediate regulation of pain-related distress of neonates was examined in two studies. The overall pooled effect showed that rocking/holding was more efficacious than a no-treatment control at improving pain-related regulation for neonates (SMD -0.75; 95% CI -1.20 to -0.30; I^2 = 0%). Please see Analysis 19.1 for the complete analysis.

ii) Older infants

a. Pain reactivity

The effect of holding an infant during a routine immunization on pain reactivity was examined in one large study. The results from this study indicate that simply holding the infant during the immunization was not efficacious in reducing pain-related distress pain reactivity than the no-treatment control (SMD 0.23; 95% CI -0.15 to 0.62).

Summary of treatment effects: holding/rocking

Neonates. The results indicate that holding was not efficacious in reducing immediate pain reactivity but substantial heterogeneity reduces our confidence in these pooled findings. Of side note, one non-randomized study excluded from analyses (Weissman 2009) found an effect for pain reactivity. However, despite the lack of effect for pain reactivity, pooled findings suggest there was sufficient evidence to suggest rocking/holding interventions for pain-related regulation.
Older infants. There is limited evidence that suggests holding is not effective intervention for pain-related distress reactivity in older infants.

6. Simulated rocking and water

(Please see 'Summary of results Table 8' for SMDs and CIs for the effect of simulated rocking and water on pain-related distress reactivity).

i) Preterm infants

a. Pain reactivity

One study assessed the efficacy of simulated rocking and water on the pain-related distress pain reactivity of preterm infants. This study showed that the intervention was no better than a no-treatment control in reducing pain reactivity in preterm infants (SMD 0.00; 95% CI -0.59 to 0.59).

Summary of treatment effects: artificial rocking

Preterm infants. Results from one study with satisfactory/marginal quality indicates that simulated rocking and water is not an efficacious intervention for reducing pain-related distress pain reactivity for preterm infants.

7. Touch/massage-related

(Please see 'Summary of results Table 9' for SMDs and CIs for the effect of touch/massage on pain-related distress reactivity and immediate regulation).

i) Preterm infants

a. Immediate regulation

Two small studies investigated the effect of touch on the immediate regulation of pain-related distress in preterm infants. The results suggest that touch/massage is not efficacious in improving the immediate regulation of preterm infants following a painful procedure (SMD -0.71; 95% CI -2.33 to 0.90). Confidence in these results is limited as the heterogeneity coefficient is substantial (I^2 = 86%) due to one study finding no effect and one study finding a significant effect. See Analysis 22.1 for the complete analysis.

ii) Neonates

a. Pain reactivity

One small study assessed the efficacy of touch/massage on pain-related distress pain reactivity of neonates. Infants received sensorial saturation which consisted of massage, voice, eye contact, and perfume during the procedure. Massage was no more efficacious than a no-treatment control (SMD -0.30; 95% CI -0.92 to 0.32).

b. Immediate regulation

One medium-sized study investigated the effect of touch on the immediate regulation of pain-related distress of neonates. This study did not find it to be efficacious in improving immediate regulation of pain-related distress in neonates (SMD -0.24; 95% CI -0.73 to 0.24).

iii) Older infants

a. Pain reactivity

Therapeutic touch in older infants was assessed in one small study and it was found to be no more efficacious than a no-treatment control in reducing pain-related distress pain reactivity (SMD -0.21; 95% CI -0.84 to 0.41).

Summary of treatment effects: touch or massage-related

Preterm infants. The findings from this review suggest that evidence does not support touch/massage-related interventions as efficacious in improving the immediate regulation of pain-related regulation for preterm infants. However, caution is warranted given the presence of substantial heterogeneity. Of side note, one study not included in the analysis due to use of a physiological outcome measure (Diego 2009) demonstrated an effect on heart rate.

Neonates. Limited evidence suggests touch/massage-related interventions are not an efficacious intervention to reduce pain reactivity or immediate regulation in neonates. Of peripheral note, one study that was not included in the analysis due to methodological choices (Cignacco 2008) provides further support that massage is not an efficacious intervention for reducing acute pain-related distress in neonates.

Older infants. Limited evidence suggests that therapeutic touch is not efficacious in reducing pain-related distress pain reactivity in older infants. It should be noted that this one study examined a type of touch that does not require physical contact.

8. Environment modification

(Please see 'Summary of results Table 10' for SMDs and CIs for the effect of environment modification on pain-related distress reactivity and immediate regulation).

i) Preterm infants

a. Pain reactivity

Two studies investigated the effect of environment modification on pain-related distress reactivity in preterm infants. Environment

modification was not found to be efficacious in reducing pain-related distress pain reactivity (SMD -6.44; 95% CI -17.13 to 4.26). Despite a similar description of techniques (e.g. attenuated light and noise, lateral posturing, and opportunities to grasp/suck), the heterogeneity quotient was substantial (I^2 = 97%), as one study found a small significant finding for the treatment and one study found a large significant finding for the treatment. See Analysis 26.1 for the complete analysis for this comparison.

b. Immediate regulation

The effect of environment modification on the immediate regulation of preterm infant pain-related distress was assessed by one small cross-over study. This intervention was efficacious at reducing the pain-related distress of preterm infants over time following the painful stimuli (SMD -4.01; 95% CI -5.26 to -2.77).

Summary

Preterm infants. While the pooled result from two studies suggest that environmental modification was not efficacious for pain reactivity, this must be interpreted with caution due to the presence of substantial heterogeneity. There is limited evidence to suggest that environmental modification may be efficacious for immediate pain-related distress regulation.

9. Toy distraction

(Please see 'Summary of results Table 11' for SMDs and CIs for the effect of toy distraction on pain-related distress reactivity and immediate regulation).

i) Older infants

a. Pain reactivity

Three studies with four treatment arms investigated the effects of toy distraction on the pain-related distress pain reactivity of older infants. One study had two separate treatment arms, both of which involved toy distraction and were included as separate studies (Hillgrove Stuart 2008). The control group was split between the two arms. Overall, toy distraction was not found to be efficacious in reducing pain-related distress pain reactivity in older infants (SMD -0.10; 95% CI -0.35 to 0.14). There was no substantial heterogeneity between the four treatment arms. Please see Analysis 28.1 .

b. Immediate regulation

The effect of toy distraction on the pain-related regulation of older infants was investigated in one moderately sized study with two treatment arms. Toy distraction was not efficacious in improving pain-related distress regulation for older infants (SMD -0.08; 95% CI -0.50 to 0.33). See Analysis 29.1 for the complete analysis.

Summary of treatment effects: toy distraction

Older infants. Sufficient evidence suggests that toy distraction is not more efficacious than a no-treatment control for reducing pain-related distress reactivity in older infants. Limited evidence also suggests that it was also not efficacious for improving pain-related regulation.

10. Video distraction

(Please see 'Summary of results Table 12' for SMDs and CIs for the effect of video distraction on pain-related distress reactivity and immediate regulation).

i) Older infants

a. Pain reactivity

One medium-sized study examined the impact of video on the pain-related distress of older infants. Infants were always distracted with a video during the intervention and toys may or may not have been used in conjunction. The results indicate that this intervention was efficacious in reducing pain-related distress pain reactivity for older infants (SMD -0.70; 95% CI -1.13 to -0.27).

b. Immediate regulation

The effects of video distraction on immediate regulation of pain-related distress for older infants were assessed in one large study. The results indicate that this intervention was more efficacious than a no-treatment control at improving pain-related distress regulation in older infants (SMD -0.84; 95% CI -1.20 to -0.47).

Summary of treatment effects: video distraction

Older infants. Results from this review indicate that there is evidence from one low/unknown quality study that video distraction is efficacious in reducing pain-related reactivity and one study of satisfactory/marginal quality that demonstrates efficacy for the immediate regulation of pain-related distress.

11. Structured parent involvement

(Please see 'Summary of results Table 13' for SMDs and CIs for the effect of structured parent involvement on pain-related distress reactivity and immediate regulation).

i) Older infants

a. Pain reactivity

The effect of structured parent involvement was investigated in three studies. Structured parent involvement included studies whereby parents were given specific instructions of what to do with their infant during a painful procedure (either verbally/verbal coaching or via information sheet). The overall pooled effect indicates that structured parent involvement was no more efficacious than a no-treatment control group in reducing pain-related distress pain reactivity in older infants (SMD -0.26; 95% CI -0.70 to 0.17; I^2 = 60%). However, when the study that had the outlying SMD was removed, the pooled effect became significant (SMD -0.49; 95% CI -0.83 to -0.14; I^2 = 0%). Please see Analysis 32.1 for the complete analysis.

b. Immediate regulation

The efficacy of structured parent involvement on the regulation of pain-related distress for older infants was assessed in one very large study. In this study, parents were instructed to stand by their infant's head and talk and touch during the painful procedure. The intervention was not more efficacious than a no-treatment control in reducing pain upon regulation for older infants (SMD 0.02; 95% CI -0.21 to 0.25).

Summary of treatment effects: structured parent involvement

i) Pain reactivity

Older infants. Taken together, the pooled results from this review suggests that structured parent involvement did not impact infant pain reactivity. However, extreme caution should be applied to this finding as post-hoc heterogeneity analyses contradicted this finding when one outlying study was removed.

ii) Pain regulation

Older infants. Taken together, the results from one study included in this review suggests that structured parent involvement did not impact infant pain regulation. However, extreme caution should be applied to this finding as it is based on one study. In addition, one excluded study suggested parental involvement significantly improved time to regulate post-immunization (Felt 2000).

12. Simulated Mother's voice

(Please see 'Summary of results Table 14' for SMDs and CIs for the effect of mother's voice on pain-related distress reactivity).

i) Preterm infants

a. Pain reactivity

Based on one small study, simulated mother's voice, which was modified to sound like it would in utero, was not efficacious at reducing pain-related distress pain reactivity for preterm infants (SMD -0.29; 95% CI -0.94 to 0.35).

Summary of treatment effects: mother's voice

Preterm infants. Results from one study of satisfactory/marginal quality indicated that simulated mother's voice was not more efficacious than a no-treatment control for reducing pain-related distress reactivity for preterm infants.

13. Parent presence

(Please see 'Summary of results Table 15' for SMDs and CIs for the effect of parent presence on pain-related distress immediate regulation).

i) Older infants

a. Immediate regulation

Parental presence alone during the painful medical procedure was not found in one trial to be efficacious in reducing pain-related distress upon immediate regulation for older infants (SMD 0.00; 95% CI -0.24 to 0.23).

Summary of treatment effects: parent presence

i) Immediate regulation

Older infants. The results from one very large study of satisfactory/marginal quality indicated that parent presence was not more efficacious than a no-treatment control for improving immediate regulation of pain-related distress for older infants.

Publication bias

In order to help minimize the influence of publication bias (i.e. the bias resulting from analyzing only published studies), we used systematic methods to obtain studies that were not published. This included contacting researchers and clinicians broadly through emails and list-serves, using dissertation search engines, contacting specific researchers who presented data at relevant pediatric pain or pain conferences we attended, and searching databases that register studies in progress. It is highly unlikely we were able to locate every relevant study but a methodical attempt was made. We also

plan to broadly disseminate our findings, both publishing parts of this review and presenting it at conferences, with an invitation to researchers and clinicians who have worked on studies currently omitted to contact us, helping to further minimize this bias for future updates.

DISCUSSION

Our goal for the review was to assess the efficacy of non-pharmacological interventions for the management of acute pain and distress in infants and children up to three years of age. Given our belief in the steep developmental trajectory that occurs during infancy and our belief that this steep trajectory influences the efficacy of non-pharmacological treatments, we conducted our analyses for preterms, neonates, and older infants separately. Moreover, due to the fact that different studies may measure acute pain at different time points and are therefore distinct, we analyzed pain reactivity (the reaction within 30 seconds after the painful stimulus) and pain-related regulation (reactions that occur after the initial pain reaction; in our review 30 seconds was chosen as the beginning of the regulation phase) data separately.

Summary of main results

The overall recommendations of this review are summarized succinctly in Table 2 with the explanation of the numbers presented in the table in the legend. These results are based on the 51 studies used for quantitative analyses.

From an age vantage point, the data suggest developmental trends of different strategies across the infant/early child stage (defined as under three years of age for this review). While sufficient evidence showed that kangaroo care was one of the most efficacious strategies for preterm infants, limited evidence suggested this was not the case for neonates, while the impact on infants older than one month is as yet unknown. In addition, there was some consistency seen across age groups whereby certain treatments had similar impacts across the ages. Non-nutritive sucking-related interventions had sufficient or limited evidence across all the age categories and pain response types, while water and touch/massage-related interventions had limited evidence suggesting they are not efficacious in any age group.

We also subdivided results according to the timing of the pain response (i.e. whether studies measured pain immediately after the acute pain stimulus or in a period that was after the immediate pain reaction). Limited evidence suggests that rocking/holding may not be efficacious for pain reactivity of neonates or older infants but sufficient evidence exists recommending it for immediate regulation of neonates (no work done in older children). This suggests that while it may not impact the initial response to pain, rocking/holding will aid in having a child soothe faster. A similar effect was potentially seen for environmental modification. Limited evidence suggests that environmental modification impacts preterm infant pain-related regulation but it has no impact on preterm infant pain reactivity. Once again, this suggests no impact on the initial response to pain but impact on how quickly the child soothes post-pain stimulus.

It should be noted that a number of treatment studies were not included (quantitatively or qualitatively) due to methodological choices (patients not randomized, using an active control group) that did not facilitate a comparison with studies using a randomized 'no-treatment' control group. The choice of using an active control group may have been an ethical one whereby researchers did not want to expose an infant to a painful stimulus without the benefit of some treatment. However, regardless of the reasons, conceptually and statistically, these studies could not be quantitatively analyzed in the review. Given some of these individual studies offer preliminary evidence that exposing an infant to a familiar odor (Goubet 2003; Goubet 2007; Rattaz 2005), feeding an infant formula (Weissman 2009), and injecting the least painful vaccine first (Ipp 2009), may result in reduced acute pain for infants, these findings warrant mention in our discussion and a call for more research. Finally, despite non-efficacious findings, the technique of positioning (e.g. prone, supine) has only been examined in one study that did not utilize a randomized controlled trial methodology (Grunau 2004). Mechanical vibration may also be a non-pharmacological technique that warrants future investigation (Baba 2010). Again, more research is warranted.

Overall completeness and applicability of evidence

An overall caveat should be noted regarding the issue of heterogeneity. Non-pharmacological interventions, by their nature, encapsulate a more diverse range of methodologies under the banner of a single heading than the common dose-response or drug comparison randomized pharmacological studies. Thus, it was not surprising that of the 35 separate comparisons run, heterogeneity was substantial in 12 cases. Heterogeneity undermines the certainty of the findings due to the presence of widely variable treatment effects among the studies being meta-analyzed. Whenever heterogeneity was an issue, we noted it in the individual summary sections. However, it should be noted that when outlier studies were removed from the analysis (deduced from looking at absolute standardized mean differences from each study, but leaving at least two studies in the analysis), in six cases the results still held up. This suggests greater confidence in the pooled result, in spite of the presence of heterogeneity. In the remaining six cases, there were five comparisons with only two studies (thus no analysis would be run if one study was dropped), thus the qualifier of 'limited evidence' was applied (with the need for more research). There was one comparison (strucutured parent involvement for older infants) where the pooled result became significant after the

heterogeneous study was removed. Thus, although not reported in Table 2, this summary table of the entire review does take into account heterogeneity.

Finally, in terms of study quality, we found that using the modified Yates scale (Yates 2005) adequately allowed us to examine the quality of non-pharmacological treatments. The majority of studies were of satisfactory/marginal quality with approximately four studies of low/unknown quality. Better reporting or methodological practices surrounding blinding (at randomization, allocation, and data collection/coding stages) and statistics (reporting power analyses and ensuring adequate sample sizes) would have raised the scores of many studies. Unfortunately, due to the scale, studies that had significant results (where power or adequate sample size was not an issue), were still penalized if they did not report power and adequacy of sample size, lowering their quality scores.

Despite these limitations, this review provides a broad and in-depth critical examination of the literature on non-pharmacological interventions for reducing pain reactivity and improving pain-related regulation for infants. To our knowledge, with 51 studies included for quantitative analysis and 20 studies mentioned qualitatively in our results section, this review is one of the most comprehensive to date for non-pharmacological acute pain management in infancy and early childhood.

AUTHORS' CONCLUSIONS

Implications for practice

Preterm infants: there was sufficient evidence to recommend kangaroo care, non-nutritive sucking-related interventions, and swaddling/tucking-related interventions as efficacious for both pain reactivity and pain-related regulation. There was limited evidence that environmental modification is potentially efficacious for the regulation of pain-related distress. It should also be noted that for sucking, allowing the preterm to suck for at least three minutes prior to the painful stimulus may be especially pain-relieving. In addition, while limited evidence suggested that touch/massage and environmental modification were not efficacious, confidence in these findings is significantly reduced due to the presence of heterogeneity.

Neonates: there was sufficient evidence to recommend non-nutritive sucking-related interventions as an efficacious treatment for acute pain reactivity and pain-related regulation. The efficacy of rocking/holding was also found to be justified by the research for pain-related regulation. Finally, despite the overall finding that kangaroo care and rocking/holding were not efficacious for neonates, the presence of substantial heterogeneity undermines our confidence in these findings. Furthermore, the question was raised that kangaroo care may be efficacious if administered to neonates for over 10 minutes prior to the painful stimulus.

Older infants: there was limited evidence for the effect of non-nutritive sucking on immediate regulation and video-mediated distraction for both pain reactivity and pain-related regulation. Substantial heterogeneity was present in the analysis on structured parent involvement and thus our confidence in the finding that it is not an efficacious treatment is limited.

Implications for research

Based on the results of this review, significant gaps in the existing literature on non-pharmacological management of acute pain in infancy can be discerned. Among the more notable gaps where there was insufficient evidence for a potentially efficacious treatment based on evidence from different age ranges are:

- kangaroo care for young infants' pain reactivity and regulation (e.g. for the two-month immunization);
- sucking-related interventions for older infants' pain reactivity; and
- swaddling or tucking-related interventions for older infants' pain reactivity and pain-related regulation (e.g. two-month immunization).

In addition, preliminary work from other studies (excluded from our overall quantitative analyses for methodological reasons) suggests that exposing an infant to a familiar odor, feeding an infant formula, and administering the least painful immunization first are promising non-pharmacological interventions that may reduce infant pain in the acute setting, but more research is needed.

In terms of further research needed, given the frequency of immunization during the first years of life, it was disheartening to realize that there was no treatment for older infants that had more than one study establishing its efficacy. Although there is a substantial evidence base for pharmacological strategies such as sucrose and topical anesthetics, it would behoove researchers to spend more resources in investigating efficacious non-pharmacological pain management for older infants.

The role of parent-mediated interventions was also a cause for concern. Over the first year of life, it has been argued that the caregiver is the most important context for the infants (Pillai Riddell 2009). Currently, studies that have attempted to formally structure parent behavior have been limited and, thus, shown to be ineffective. It should be stressed that more work on better types of parent interventions, especially ones that capitalize on an infant's primary developmental need for proximity to the parent, as theorized by the widely accepted attachment theory (Bowlby 1969/1982), is needed. Teaching parents to better meet an infant's attachment needs during times of pain may lead to more efficacious interventions.

Finally, from a measurement perspective, two points were noted. One, whenever possible investigators should measure beyond the

pain reactivity phase (e.g. first 30 seconds after a painful stimulus). Treatments that may not impact initial pain reactivity but significantly impact pain-related regulation (e.g. pain reactivity greater than 30 seconds after a painful stimulus) may be unnecessarily cast aside by researchers who only focus on the immediate reactions post-immunization (see rocking/holding or environmental modification). Secondly, reporting measurements that allow meta-analytic techniques are crucial. Considerable time was spent having to interpolate means and standard deviations from studies that did not meet basic standards of statistical reporting.

ACKNOWLEDGEMENTS

We would like to thank the team at Cochrane for their patience (Jessica Thomas, Dr. Amanda C de C Williams, and Dr. John Plummer) and support with our complex search strategies (Elizabeth Ulryk, Sylvia Bickley, and Caroline Struthers) over the years. The breadth of the literature reviewed is due to their willingness to listen to our concerns and skillfully adapt the search. We would also like to thank Dr. Ken Craig and Dr. Amanda C de C Williams for their help at the protocol stage and especially for recommending the use of the Yates Scales. Finally, we would like to thank our research assistants at the OUCH lab, Diana Lisi and Jennifer Lyons, for their extensive help with all the project and data management of this review. We would also like to thank Janet Yamada and Grace Lee for helping with extraction.

In addition, we would like to thank our sources of funding over the creation of this analysis: The Canadian Institutes of Health Research has provided salary support to Dr. Pillai Riddell (Postdoctoral Fellowship, New Investigator Award) and doctoral fellowships to Ms. Laila Din Osmun and Ms. Jessica Hillgrove Stuart. The Ontario Graduate Scholarship Program has provided support to Ms. Nicole Racine, Ms. Rachel Horton, and Ms. Sara Ahola Kohut. The Social Sciences and Humanities Research Council has provided support to Ms. Nicole Racine. Fonds de Recherche en Sante du Quebec provided doctoral funding to Dr. Lindsay Uman. In addition, during this review, The Pain in Child Health (Canadian Institutes for Health Research) strategic training initiative provided support to Dr. Uman, Ms. Horton, Ms. Din Osmun, Ms. Ahola Kohut, Ms. Hillgrove-Stuart, and Ms. Racine. Finally, the Lillian Wright Foundation for Maternal-Child Health has also provided funds to: Ms. Nicole Racine, Ms. Laila Din Osmun, and Ms. Rachel Horton. In addition, operating funds for this project were provided by start up funds from York University to Dr. Pillai Riddell.

REFERENCES

References to studies included in this review

Akcan 2009 *{published data only}*
Akcan E, Yigit R, Atici A. The effect of kangaroo care on pain in premature infants during invasive procedures. *The Turkish Journal of Pediatrics* 2009;**51**:14–8.

Allen 1996 *{published data only}*
Allen KD, White DD, Walburn JN. Sucrose as an analgesic agent for infants during immunization injections. *Archives of Pediatrics & Adolescent Medicine* 1996;**150**(3):270–4.

Axelin 2006 *{published data only}*
Axelin A, Salantera S, Lehtonen L. Facilitated tucking by parents in pain management of preterm infants - a randomized crossover trial. *Early Human Development* 2006;**82**:241–7.

Axelin 2009 *{published data only}*
Axelin A, Salantera S, Kirjavainen J, Lehtonen L. Oral glucose and parental holding preferable to opioid in pain management in preterm infants. *Clinical Journal of Pain* 2009;**25**:138–45.

Bauchner 1996 *{published data only}*
Bauchner H, Vinci R, Bak S, Pearson C, Corwin MJ. Parents and procedures: a randomized controlled trial. *Pediatrics* 1996;**98**(5):861–7.

Bellieni 2001 *{published data only}*
Bellieni CV, Buonocore G, Nenci A, Franci N, Cordelli DM, Bagnoli F. Sensorial saturation: an effective analgesic tool for heel-prick in preterm infants. *Biology of the Neonate* 2001;**80**(1):15–8.

Bellieni 2002 *{published data only}*
Bellieni CV, Bagnoli F, Perrone S, Nenci A, Cordelli DM, Fusi M, et al.Effect of multisensory stimulation on analgesia in term neonates: a randomized controlled trial. *Pediatric Research* 2002;**51**(4):460–3.

Blass 1999 *{published data only}*
Blass EM, Watt LB. Suckling- and sucrose-induced analgesia in human newborns. *Pain* 1999;**83**(3):611–23.

Bo 2000 *{published data only}*
Bo LK, Callaghan P. Soothing pain-elicited distress in Chinese neonates. *Pediatrics* 2000;**105**(4):E49. [DOI: 10.1542/peds.105.4.e49]

Bustos 2008 *{published data only}*
Bustos T, Jaaniste T, Salmon K, Champion GD. Evaluation of a brief parent intervention teaching coping-promoting behavior for the infant immunization context: a randomized controlled trial. *Behavior Modification* 2008;**32**(4):450–67. [DOI: 10.1177/0145445507309031]

Campos 1994 *{published data only}*
Campos RG. Rocking and pacifiers: two comforting interventions for heelstick pain. *Research in Nursing &*

Health 1994;**17**(5):321–31.

Carbajal 1999 ***{published data only}***
Carbajal R, Chauvet X, Couderc S, Olivier-Martin M. Randomised trial of analgesic effects of sucrose, glucose, and pacifiers in term neonates. *BMJ* 1999;**319**(7222):1393–7.

Carbajal 2003 ***{published data only}***
Carbajal R, Veerapen S, Couderc S, Jugie M, Ville Y. Analgesic effect of breast feeding in term neonates: Randomised controlled trial. *British Medical Journal (Clinical Research Ed.)* 2003;**326**(7379):13.

Castral 2008 ***{published data only}***
Castral TC, Warnock F, Leite AM, Haas VJ, Scochi, CGS. The effects of skin-to-skin contact during acute pain in preterm newborns. *European Journal of Pain* 2008;**12**: 464–471.

Catelin 2005 ***{published data only}***
Catelin C, Tordjman S, Morin V, Oger E, Sizun J. Clinical, physiologic, and biologic impact of environmental and behavioral interventions in neonates during a routine nursing procedure. *Journal of Pain* 2005;**6**(12):791–7. [DOI: 10.1016/j.jpain.2005.07.010]

Chermont 2009 ***{published data only}***
Chermont AG, Falcao LFM, de Souza Silva EHL, de Cassia Xavier Balda R, Guinsburg R. Skin-to-skin contact and/or oral 25% dextrose for procedural pain relief for term newborn infants. *Pediatrics* 2009;**124**:1101–7. [DOI: 10.1542/peds.2009-0993]

Cohen 2002 ***{published data only}***
Cohen LL. Reducing infant immunization distress through distraction. *Health Psychology* 2002;**21**(2):207–11.

Cohen 2006 ***{published data only}***
Cohen LL, MacLaren JE, Fortson BL, Friedman A, DeMore M, Lim CS, et al.Randomized clinical trial of distraction for infant immunization pain. *Pain* 2006;**125**(1-2):165–71.

Comaru 2009 ***{published data only}***
Comaru T, Miura E. Postural support improves distress and pain during diaper change in preterm infants. *Journal of Perinatology* 2009;**29**:504–7. [DOI: 10.1038/jp.2009.13]

Corbo 2000 ***{published data only}***
Corbo MG, Mansi G, Stagni A, Romano A, van den Heuvel J, Capasso L et al. Nonnutritive sucking during heelstick procedures decreases behavioral distress in the newborn infant. *Biology of the Neonate* 2000;**77**(3):162–7.

Corff 1995 ***{published data only}***
Corff KE, Seideman R, Venkataraman PS, Lutes L, Yates B. Facilitated tucking: a nonpharmacologic comfort measure for pain in preterm neonates. *Journal of Obstetric, Gynecologic, & Neonatal Nursing* 1995;**24**(2):143–7.

Cramer-Berness 2005 ***{published data only}***
Cramer-Berness LJ, Friedman AG. Behavioral interventions for infant immunizations. *Children's Health Care* 2005;**34** (2):95–111.

Cramer-Berness 2005b ***{published data only}***
Cramer-Berness LJ. A comparison of behavioral interventions for infant immunizations. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 2005;**66**(12-B):6916.

Curtis 2007 ***{published data only}***
Curtis SJ, Jou H, Ali S, Vandermeer B, Klassen T. A randomized controlled trial of sucrose and/or pacifier as analgesia for infants receiving venipuncture in a pediatric emergency department. *BMC Pediatrics* 2007;**7**:27.

de Sousa 2008 ***{published data only}***
de Sousa Freire NB, Santos Garcia JB, Carvalho Lamy Z. Evaluation of analgesic effect of skin-to-skin contact compared to oral glucose in preterm neonates. *Pain* 2008; **139**(1):28–33. [DOI: 10.1016/j.pain.2008.02.031]

Elserafy 2009 ***{published data only}***
Elserafy FA, Alsaedi SA, Louwrens J, Bin Sadiq B, Mersal AY. Oral sucrose and a pacifier for pain relief during simple procedures in preterm infants: a randomized controlled trial. *Annals of Saudi Medicine* 2009;**29**(3):184–8. [DOI: 10.4103/0256-4947.52821]

Fearon 1997 ***{published data only}***
Fearon I, Kisilevsky BS, Hains SMJ, Muir DW, Tranmer J. Swaddling after heel lance: age-specific effects on behavioral recovery in preterm infants. *Developmental and Behavioral Pediatrics* 1997;**18**(4):222–32.

Ferber 2008 ***{published data only}***
Ferber SG, Makhoul IR. Neurobehavioural assessment of skin-to-skin effects on reaction to pain in preterm infants: a randomized, controlled within-subject trial. *Acta Paediatrica* 2008;**97**(2):171–6. [DOI: 10.1111/j.1651-2227.2007.00607.x]

Gormally 2001 ***{published data only}***
Gormally S, Barr RG, Wertheim L, Alkawaf R, Calinoiu N, Young SN. Contact and nutrient caregiving effects on newborn infant pain responses. *Developmental Medicine & Child Neurology* 2001;**43**(1):28–38.

Gray 2000 ***{published data only}***
Gray L, Watt L, Blass EM. Skin-to-skin contact is analgesic in healthy newborns. *Pediatrics* 2000;**105**(1):e14.

Greenberg 2002 ***{published data only}***
Greenberg CS. A sugar-coated pacifier reduces procedural pain in newborns. *Pediatric Nursing* 2002;**28**(3):271–7.

Herrington 2007 ***{unpublished data only}***
Herrington CJ. Reducing pain of heelstick in premature infants with gentle human touch. Wayne State University Doctoral Dissertation unpublished:1–229.

Hill 2005 ***{published data only}***
Hill S, Engle S, Jorgensen J, Kralik A, Whitman K. Effects of facilitated tucking during routine care of infants born preterm. *Pediatric Physical Therapy* 2005;**17**(2):158–63.

Hillgrove Stuart 2008 ***{unpublished data only}***
Hillgrove Stuart J, Pillai Riddell R. Distraction as a pain management strategy for infants. Master's Thesis unpublished.

Im 2008 *{published data only}*
Im H, Kim E, Park E, Sung K, Oh W. Pain reduction of heel stick in neonates: Yakson compared to non-nutritive sucking. *Journal of Tropical Pediatrics* 2008;**54**(1):31–5.

Ipp 2004 *{published data only}*
Ipp M, Taddio A, Goldbach M, Ben David S, Stevens B, Koren G. Effects of age, gender and holding on pain response during infant immunization. *Canadian Journal of Clinical Pharmacology* 2004;**11**(1):e2–7.

Jain 2006 *{published data only}*
Jain S, Kumar P, McMillan DD. Prior leg massage decreases pain responses to heel stick in preterm babies. *Journal of Paediatrics & Child Health* 2006;**42**(9):505–8. [DOI: 10.111/j.1440-1754.2006.00912.x]

Johnston 1997 *{published data only}*
Johnston CC, Stremler RL, Stevens BJ, Horton LJ. Effectiveness of oral sucrose and simulated rocking on pain response in preterm neonates. *Pain* 1997;**72**(1-2):193–9.

Johnston 2003 *{published data only}*
Johnston CC, Stevens B, Pinelli J, Gibbins S, Filion F, Jack A, et al.Kangaroo care is effective in diminishing pain response in preterm neonates. *Archives of Pediatrics & Adolescent Medicine* 2003;**157**(11):1084–8.

Johnston 2007 *{published data only}*
Johnston CC, Filion F, Nuyt AM. Recorded maternal voice for preterm neonates undergoing heel lance. *Advances in Neonatal Care* 2007;**7**(5):258–66.

Kashaninia 2008 *{published data only}*
Kashaninia Z, Sajedi F, Rahgozar M, Noghabi FA. The effect of kangaroo care on behavioral responses to pain of an intramuscular injection in neonates. *Journal for Specialists in Pediatric Nursing* 2008;**13**(4):275–80.

Kostandy 2008 *{published data only}*
Kostandy RR, Ludington-Hoe SM, Cong X, Abouelfettoh A, Bronson C, Stankus A, et al.Kangaroo care (skin contact) reduces crying response to pain in preterm neonates: pilot results. *Pain Management Nursing* 2008;**9**(2):55–65.

Kozub 2001 *{unpublished data only}*
Kozub ML. *The effect of therapeutic touch on pain response in infants receiving injections*. Ohio: Medical College of Ohio, 2001.

Liaw 2010 *{published data only}*
Liaw J, Yang L, Ti Y, Tucker Blackburn S, Chang Y, Sun L. Non-nutritive sucking relieves pain for preterm infants during heel stick procedures in Taiwan. *Journal of Clinical Nursing* 2010;**19**:2741–51.

Liu 2010 *{published data only}*
Liu M, Lin K, Chou Y, Lee T. Using non-nutritive sucking and oral glucose solution with neonates to relieve pain: a randomised controlled trial. *Journal of Clinical Nursing* 2010;**19**:1604–11.

Morrow 2010 *{published data only}*
Morrow C, Hidinger A, Wilkinson-Faulk D. Reducing neonatal pain during routine heel lance procedures. *American Journal of Maternal/Child Nursing* 2010;**35**(6): 346–54.

Sizun 2002 *{published data only}*
Sizun J, Ansquer H, Browne J, Tordjman S, Morin JF. Developmental care decreases physiologic and behavioral pain expression in preterm neonates. *Journal of Pain* 2002;**3** (6):446–50.

Stevens 1999 *{published data only}*
Stevens B, Johnston C, Franck L, Petryshen P, Jack A, Foster G. The efficacy of developmentally sensitive interventions and sucrose for relieving procedural pain in very low birth weight neonates. *Nursing Research* 1999;**48**(1):35–43.

Ward-Larson 2004 *{published data only}*
Ward-Larson C, Horn RA, Gosnell F. The efficacy of facilitated tucking for relieving procedural pain of endotracheal suctioning in very low birthweight infants. *MCN: The American Journal of Maternal Child Nursing* 2004;**29**(3):151–6.

Whipple 2004 *{published data only}*
Whipple J. *The effect of music-reinforced nonnutritive sucking on state of preterm, low birthweight infants experiencing heelstick*. Florida: The Florida State University, 2004.

Yilmaz 2010 *{published data only}*
Yilmaz F, Arikan D. The effects of various interventions to newborns on pain and duration of crying. *Journal of Clinical Nursing* 2010;**20**:1008–17.

References to studies excluded from this review

Abedin 2008 *{published data only}*
Abedin S, Kapoor G. Peripherally inserted central venous catheters are a good option for prolonged venous access in children with cancer. *Pediatric Blood Cancer* 2008;**51**: 251–5.

Aguirre 2008 *{published data only}*
Aguirre Unceta-Barrenechea A, Saitua Iturriaga G, Sainz de Rozas Aparicio I, Riveira Fernandez D. Analgesia when taking heel-lance blood in the newborn. *Anales de Pediatria* 2008;**69**(6):544–7.

Al-Bekaa 2003 *{published data only}*
Al-Bekaa S. Harmonic scalpel tonsillectomy vs diathermy tonsillectomy: a comparative study. *Australian Journal of Otolaryngology* 2003;**6**(2):78–80.

Aslanabadi 2008 *{published data only}*
Aslanabadi S, Ghalehgolab-Behbahan A, Zarrintan S, Jamshidi M, Seyyedhejazi M. Transanal one-stage endorectal pull-through for Hirschsprung's disease: a comparison with the staged procedures. *Pediatric Surgery International* 2008; **24**:925–9.

Axelin 2010 *{published data only}*
Axelin A, Lehtonen L, Pelander T, Slantera S. Mothers' different styles of involvement in preterm infant pain care. *Journal of Obstetric, Gynecologic and Neonatal Nursing* 2010; **39**:415–24.

Bellieni 2003 *{published data only}*
Bellieni CV, Bagnoli F, Cordelli DM, Perrone S, Nenci A, Buonocore G. Analgesia for premature children. *Italian Journal of Pediatrics* 2003;**29**(4):286–90.

Bellieni 2007 *{published data only}*
Bellieni CV, Cordelli DM, Marchi S, Ceccarelli S, Perrone S, Maffei M, et al.Sensorial saturation for neonatal analgesia. *Clinical Journal of Pain* 2007;**23**(3):219–21.

Boots 2010 *{published data only}*
Boots BK, Edmundson EE. A controlled, randomised trial comparing single to multiple application lidocaine analgesia in paediatric patients undergoing urethral catheterisation procedures. *Journal of Clinical Nursing* 2010;**19**:744–8.

Bueno 2010 *{unpublished data only}*
Bueno M, Bussotti E, DaSilva A, Leao E. Non nutritive sucking and swaddling for pain relief in term neonates. Abstract from the International Symposium on Pediatric Pain 2010:1.

Campbell-Yeo 2009 *{published data only}*
Campbell-Yeo ML, Johnston C, Joseph KS, Feeley NL, Chambers CT, Barrington KJ. Co-bedding as a comfort measure for twins undergoing painful procedures (CComForT Trial). *BMC Pediatrics* 2009;**9**:76. [DOI: 10.1186/1471-2431-9-76]

Campos 1989 *{published data only}*
Campos RG. Soothing pain-elicited distress in infants with swaddling and pacifiers. *Child Development* 1989;**60**(4):781–92.

Carlson 2000 *{published data only}*
Carlson KL, Broome M, Vessey JA. Using distraction to reduce reported pain, fear, and behavioral distress in children and adolescents: a multisite study. *Journal of the Society of Pediatric Nurses* 2000;**5**(2):75–85.

Cignacco 2008 *{published data only}*
Cignacco E, Hamers JP, Van Lingen RA, Zimmermann LJI, Muller R, Gessler P, et al.Pain relief in ventilated preterm infants during endotracheal suctioning: a randomized controlled trial. *Swiss Medical Weekly* 2008;**138**(43-44):635–45.

Cologna 1999 *{published data only}*
Cologna M, Sperandio L. The effect of two different methods of heel lancing on pain reaction in preterm neonates. *Assistenza Infermieristica e Ricerca:Air* 1999;**18**(4):185–92.

Cong 2009 *{published data only}*
Cong X, Ludington-Hoe SM, McCain G, Fu P. Kangaroo care modifies preterm infant heart rate variability in response to heel stick pain: pilot study. *Early Human Development* 2009;**85**:561–7. [DOI: 10.1016/j.earlhumdev.2009.05.012]

D'Agostino 2008 *{published data only}*
D'Agostino R, Tarantino V, Calevo MG. Blunt dissection versus electronic molecular resonance bipolar dissection for tonsillectomy: operative time and intraoperative and postoperative bleeding and pain. *International Journal of Pediatric Otorhinolaryngology* 2008;**72**(7):1077–84.

Dahlquist 2002 *{published data only}*
Dahlquist LM, Pendley JS, Landthrip DS, Jones CL, Steuber CP. Distraction intervention for preschoolers undergoing intramuscular injections and subcutaneous port access. *Health Psychology* 2002;**21**(1):94–9.

Diego 2009 *{published data only}*
Diego MA, Field T, Hernandez-Reif M. Procedural pain heart rate responses in massaged preterm infants. *Infant Behavior & Development* 2009;**32**:226–9. [DOI: 10.1016/j.infbeh.2008.12.001]

Dilen 2010 *{published data only}*
Dilen B, Elseviers M. Oral glucose solution as pain relief in newborns: results of a clinical trial. *Birth: Issues in Perinatal Care* 2010;**37**(2):98–105.

Drago 2009 *{published data only}*
Drago LA, Singh SB, Douglass-Bright A, Yaa Yiadom M, Baumann BM. Efficacy of ShotBlocker in reducing pediatric pain associated with intramuscular injections. *American Journal of Emergency Medicine* 2009;**27**:536–43. [DOI: 10.1016/j.ajem.2008.04.011]

Duncan 2004 *{published data only}*
Duncan B, Barton L, Edmonds D, Blashill BM. Parental perceptions of the therapeutic effect from osteopathic manipulation or acupuncture in children with spastic cerebral palsy. *Clinical Pediatrics* 2004;**43**(4):349–53.

El-Naggar 2010 *{published data only}*
El-Naggar W, Yiu A, Mohamed A, Shah V, Manley J, McNamara P, et al.Comparison of pain during two methods of urine collection in preterm infants. *Pediatrics* 2010;**125**(6):1224–9.

Favara-Scacco 2001 *{published data only}*
Favara-Scacco C, Smirne G, Schiliro G, Di Cataldo A. Art therapy as support for children with leukemia during painful procedures. *Medical & Pediatric Oncology* 2001;**36**(4):474–80.

Felt 2000 *{published data only}*
Felt BT, Mollen E, Diaz S, Renaud E, Zeglis M, Wheatcroft G, Mendelow D. Behavioral interventions reduce infant distress at immunization. *Archives of Pediatric and Adolescent Medine* 2000;**154**:719–724.

Gedaly-Duff 1992 *{published data only}*
Gedaly-Duff V, Burns C. Reducing children's pain-distress associated with injections using cold: a pilot study. *Journal of the American Academy of Nurse Practitioners* 1992;**4**(3):95–9.

Gold 2006 *{published data only}*
Gold JI, Kim SH, Kant AJ, Joseph MH, Rizzo AS. Effectiveness of virtual reality for pediatric pain distraction during i.v. placement. *Cyberpsychology & Behavior* 2006;**9**(2):207–12.

Gonzalez 1989 *{published data only}*
Gonzalez JC, Routh DK, Saab PG, Armstrong FD, Shifman L, Guerra E et al. Effects of parent presence on children's reactions to injections: behavioral, physiological, and subjective aspects. *Journal of Pediatric Psychology* 1989;**14**(3):449–62.

Gonzalez 1993 *{published data only}*
Gonzalez JC, Routh DK, Armstrong FD. Effects of maternal distraction versus reassurance on children's reactions to injections. *Journal of Pediatric Psychology* 1993;**18**(5): 593–604.

Goubet 2003 *{published data only}*
Goubet N, Rattaz C, Pierrat V, Bullinger A, Lequien P. Olfactory experience mediates response to pain in preterm newborns. *Developmental Psychobiology* 2003;**42**(2): 171–80.

Goubet 2007 *{published data only}*
Goubet N, Strasbaugh K, Chesney J. Familiarity breeds content? Soothing effect of a familiar odor on full-term newborns. *Journal of Developmental & Behavioral Pediatrics* 2007;**28**(3):189–94.

Greenberg 1997 *{published data only}*
Greenberg CS. *An evaluation of the efficacy of selected nonpharmacologic pain interventions in infants*. San Diego: University of San Diego, 1997.

Grunau 2004 *{published data only}*
Grunau RE, Linhares MB, Holsti L, Oberlander TF, Whitfield MF. Does prone or supine position influence pain responses in preterm infants at 32 weeks gestational age?. *Clinical Journal of Pain* 2004;**20**(2):76–82.

Hanson 2010 *{published data only}*
Hanson D, Hall W, Mills LL, Au S, Bhagat R, Hernandez M, et al.Comparison of distress and pain in infants randomized to groups receiving standard versus multiple immunizations. *Infant Behavior and Development* 2010;**33**: 289–96.

Harrison 2000 *{published data only}*
Harrison LL, Williams AK, Berbaum ML, Stem JT, Leeper J. Physiologic and behavioral effects of gentle human touch on preterm infants. *Research in Nursing & Health* 2000;**23** (6):435–46.

Hatem 2006 *{published data only}*
Hatem TP, Lira PI, Mattos SS. The therapeutic effects of music in children following cardiac surgery. *Jornal de Pediatria* 2006;**82**(3):186–92.

He 2010 *{published data only}*
He H, Jahja R, Lee T, Ang E, Sinnappan R, Vehvilainen-Julkunen K, et al.Nurses' use of non-pharmacological methods in children's postoperative pain management: educational intervention study. *Journal of Advanced Nursing* 2010;**66**(11):2398–409.

Heden 2009 *{published data only}*
Heden L, von Essen L, Ljungman G. Randomized interventions for needle procedures in children with cancer. *European Journal of Cancer Care* 2009;**18**:358–63. [DOI: 10.1111/j.1365-2354.2008.00939.x]

Hoffman 2006 *{published data only}*
Hoffman HG, Richards TL, Bills AR, Van Oostrom T, Magula J, Seibel EJ, et al.Using fMRI to study the neural correlates of virtual reality analgesia. *CNS Spectrums* 2006; **11**(1):45–51.

Holsti 2005 *{published data only}*
Holsti L, Grunau RE, Oberlander TF, Whitfield MF. Prior pain induces heightened motor responses during clustered care in preterm infants in the NICU. *Early Human Development* 2005;**81**(3):293–302.

Hsu 1995 *{published data only}*
Hsu CY, Huang LM, Lee CY, Lin TY, Lee PI, Chen JM. Local massage after vaccination enhances the immunogenicity of diphtheria- tetanus-pertussis vaccine. *Pediatric Infectious Disease Journal* 1995;**14**(7):567–72.

Huang 1999 *{published data only}*
Huang FY, Huang LM. Effect of local massage on vaccination: DTP and DTPa. *Acta Paediatrica Taiwanica* 1999;**40**(3):166–70.

Huang 2004 *{published data only}*
Huang CM, Tung WS, Kuo LL, Ying-Ju C. Comparison of pain responses of premature infants to the heelstick between containment and swaddling. *Journal of Nursing Research* 2004;**12**(1):31–40.

Ida 2008 *{published data only}*
Ida S, Yoshimura N, Nakacho M, Ota M, Mine H, Moriyama H, et al.Evaluation of the use of PenNeedle 32G Taper in children being treated with growth hormone. *Clinical Pediatric Endocrinology* 2008;**17**(1):1–7.

Ipp 2007 *{published data only}*
Ipp M, Taddio A, Sam J, Goldbach M, Parkin PC. Vaccine-related pain: randomised controlled trial of two injection techniques. *Archives of Disease in Childhood* 2007;**92**(12): 1105–8.

Ipp 2009 *{published data only}*
Ipp M, Parkin PC, Lear N, Goldbach M, Taddio A. Order of vaccine injection and infant pain response. *Archives of Pediatrics & Adolescent Medicine* 2009;**163**(5):469–72.

Jackson 2008 *{published data only}*
Jackson LA, Starkovich P, Dunstan M, Yu O, Nelson J, Dunn J, et al.Prospective assessment of the effect of needle length and injection site on the risk of local reactions to the fifth diphtheria-tetanus-acellular pertussis vaccination. *Pediatrics* 2008;**21**(3):e646–52.

Jo 2007 *{published data only}*
Jo SH, Mathiasen RA, Gurushanthaiah D. Prospective, randomized, controlled trial of a hemostatic sealant in children undergoing adenotonsillectomy. *Otolaryngology - Head & Neck Surgery* 2007;**137**(3):154–8.

Johnston 2007b *{published data only}*
Johnston CC, Gagnon A, Rennick J, Rosmus C, Patenaude H, Ellis J, et al.One-on-one coaching to improve pain assessment and management practices of pediatric nurses. *Pediatric Nursing* 2007;**22**(6):467–78.

Johnston 2008a *{published data only}*
Johnston CC, Filion F, Campbell-Yeo M, Goulet C, Bell L, McNaughton K, et al.Kangaroo mother care diminishes pain from heel lance in very preterm neonates: a crossover trial. *BMC Pediatrics* 2008;**8**:13. [DOI: 10.1186/1471-2431-8-13]

Johnston 2009 *{published data only}*
Johnston CC, Filion F, Campbell-Yeo M, Goulet C, Bell L, McNaughton K, et al.Enhanced kangaroo mother care for heel lance in preterm neonates: a crossover trial. *Journal of Perinatology* 2009;**29**:51–6. [DOI: 10.1038/jp.2008.113]

Kivijärvi 2008 *{published data only}*
Kivijärvi R, Palviainen J. "Just like I had kicked my toe" -- Children's experiences of postsurgical pain. *Sairaanhoitaja* 2008;**81**(1):27–9.

Koivusalo 2009 *{published data only}*
Koivusalo AI, Korpela R, Wirtavuori K, Piiparinen S, Rintala RJ, Pakarinen MP. A single-blinded, randomized comparison of laparoscopic versus open hernia repair in children. *Pediatrics* 2009;**123**(1):332–7.

Leclair 2007 *{published data only}*
Leclair MD, Plattner V, Mirallie E, Lejus C, Nguyen JM, Podevin G, et al.Laparoscopic pyloromyotomy for hypertrophic pyloric stenosis: a prospective, randomized controlled trial. *Journal of Pediatric Surgery* 2007;**42**(4): 692–8.

Li 2007 *{published data only}*
Li HCW, Lopez V, Lee TLI. Effects of preoperative therapeutic play on outcomes of school-age children undergoing day surgery. *Research in Nursing & Health* 2007; **30**(3):320–32.

Ludington-Hoe 2005 *{published data only}*
Ludington-Hoe SM, Hosseini R, Torowicz DL. Skin-to-skin contact (kangaroo care) analgesia for preterm infant heel stick. *AACN Clinical Issues* 2005;**16**(3):373–87.

MacLaren 2005 *{published data only}*
MacLaren JE, Cohen LL. A comparison of distraction strategies for venipuncture distress in children. *Journal of Pediatric Psychology* 2005;**30**(5):387–96.

Manne 1990 *{published data only}*
Manne SL, Redd WH, Jacobsen PB, Gorfinkle K, Schorr O, Rapkin B. Behavioral intervention to reduce child and parent distress during venipuncture. *Journal of Consulting & Clinical Psychology* 1990;**58**(5):565–72.

Marchisotti 2007 *{published data only}*
Marchisotti FG, Carvalho LR, Berger K, Arnhold IJ, Mendonca BB. Growth hormone (GH) deficiency treatment in children: comparison between uses of pen versus bottles/ syringes on GH administration. *Arquivos Brasileiros de Endocrinologia e Metabologia* 2007;**51**(7):1093–6.

Marec-Berard 2009 *{published data only}*
Marec-Berard P, Bissery A, Kebaili K, Schell M, Aubert F, Gaillard S, et al.A positioning pillow to improve lumbar puncture success rate in paediatric haematology-oncology patients: a randomized controlled trial. *BMC Cancer* 2009; **9**:21. [DOI: 10.1186/1471-2407-9-21]

Marin Gabriel 2010 *{published data only}*
Marin Gabriel MA, Llana Martin I, Lopez Escobar A, Fernandez Villalba E, Romero Blanco I, Touza Pol P. Randomized controlled trial of early skin-to-skin contact: effects on the mother and the newborn. *Acta Paediatrica* 2010;**99**:1630–4.

McCarthy 2010 *{published data only}*
McCarthy A, Lieiber C, Hanrahan K, Zimmerman B, Westhus N, Allen S. Impact of parent-provided distraction on child responses to an IV insertion. *Children's Health Care* 2010;**39**:125–41.

Michel 2008 *{published data only}*
Michel RG, Weinstock BI, Tsau K. Safety and efficacy of pressure-assisted tissue-welding tonsillectomy: a preliminary evaluation. *Ear, Nose & Throat Journal* 2008;**87**(2):100–12.

Morelius 2009 *{published data only}*
Morelius E, Theodorsson E, Nelson N. Stress at three-month immunization: parents' and infants' salivary cortisol response in relation to the use of pacifier and oral glucose. *European Journal of Pain* 2009;**13**(2):202–8.

Mucignat 2004 *{published data only}*
Mucignat V, Ducrocq S, Lebas F, Mochel F, Baudon JJ, Gold F. Analgesic effects of Emla cream and saccharose solution for subcutaneous injections in preterm newborns: a prospective study of 265 injections. *Archives de Pediatrie* 2004;**11**(8):921–5.

Okan 2010 *{published data only}*
Okan F, Ozdil A, Bulbul A, Yapici Z, Nuhoglu A. Analgesic effects of skin-to-skin contact and breastfeeding in procedural pain in healthy term neonates. *Annals of Tropical Paediatrics* 2010;**30**:119–28.

Ozdogan 2010 *{published data only}*
Ozdogan T, Akman I, Cebeci D, Bilgen H, Ozek E. Comparison of two doses of breast milk and sucrose during neonatal heel prick. *Pediatrics International* 2010;**52**:175–9.

Park 2006 *{published data only}*
Park ES, Sung KS, Oh WO, Im HS, Kim ES, Kim YA, et al.Pain relieving effect of Yakson therapy for infants. *Daehan Ganho Haghoeji* 2006;**36**(6):897–904.

Phipps 2005 *{published data only}*
Phipps S, Dunavant M, Gray E, Rai SN. Massage therapy in children undergoing hematopoietic stem cell transplantation: results of a pilot trial. *Journal of Cancer Integrative Medicine* 2005;**3**(2):62–70.

Rattaz 2005 *{published data only}*
Rattaz C, Goubet N, Bullinger A. The calming effect of a familiar odor on full-term newborns. *Journal of Developmental & Behavioral Pediatrics* 2005;**26**(2):86–92.

Reichel 2007 *{published data only}*
Reichel O, Mayr D, Winterhoff J, de la Chaux R, Hagedorn H, Berghaus A. Tonsillotomy or tonsillectomy? A prospective study comparing histological and immunological findings in recurrent tonsillitis and tonsillar hyperplasia. *European Archives of Oto-Rhino-Laryngology* 2007;**264**:277–84.

Salmon 2006 *{published data only}*
Salmon K, McGuigan F, Pereira JK. Brief report: optimizing children's memory and management of an invasive medical procedure: the influence of procedural narration and

distraction. *Journal of Pediatric Psychology* 2006;**31**(5): 522–7.

Shapiro 2007 *{published data only}*
Shapiro NL, Bhattacharyya N. Cold dissection versus coblation-assisted adenotonsillectomy in children. *Laryngoscope* 2007;**117**(3):406–10.

Slater 2010 *{published data only}*
Slater R, Cornelissen L, Fabrizi L, Patten D, Yaxen J, Warley A, et al.Oral sucrose as an analgesic drug for procedural pain in newborn infants: a randomised controlled trial. *Lancet* 2010;**376**:1225–32.

Slifer 2009 *{published data only}*
Slifer KJ, DeMore M, Vona-Messersmith N, Pulbrook-Vetter V, Beck M, Dalhquist L, et al.Comparison of two brief parent-training interventions for child distress during parent-administered needle procedures. *Children's Health Care* 2009;**38**:23–48. [DOI: 10.1080/02739610802615831]

Sparks 2007 *{published data only}*
Sparks LA, Setlik J, Luhman J. Parental holding and positioning to decrease IV distress in young children: a randomized controlled trial. *Journal of Pediatric Nursing* 2007;**22**(6):440–7.

Sundararajan 2007 *{published data only}*
Sundararajan L, Parikh DH. Evolving experience with video-assisted thoracic surgery in congenital cystic lung lesions in a British pediatric center. *Journal of Pediatric Surgery* 2007;**42**:1243–50.

Tanabe 2002 *{published data only}*
Tanabe P, Ferket K, Thomas R, Paice J, Marcantonio R. The effect of standard care, ibuprofen, and distraction on pain relief and patient satisfaction in children with musculoskeletal trauma. *Journal of Emergency Nursing* 2002; **28**(2):118–25.

Vignochi 2010 *{published data only}*
Vignochi C, Teixeira PP, Nader SS. Effect of aquatic physical therapy on pain and state of sleep and wakefulness among stable preterm newborns in neonatal intensive care units [Efeitos da fisioterapia aquatica na dor e no estado de sono e vigilia de recem–nascidos pre–termo estaveis internados em unidade de terapia intensiva neonatal]. *Revista Brasileira de Fisioterapia (Brazilian Journal of Physical Therapy)* 2010;**14** (3):214–20.

Vivancos 2010 *{published data only}*
Vivancos RBZ, Leite AM, Scochi CGS, dos Santos CB. The skin to skin contact at birth and newborn crying during vaccination against hepatitis B [O contato pele a pele ao nascimento e o choro de recem–nascidos durante vacinacao contra Hepatite B]. *Acta Paul Enferm* 2010;**23**(4):461–5.

Weissman 2009 *{published data only}*
Weissman A, Aranovitch M, Blazer S, Zimmer EZ. Heel-lancing in newborns: behavioral and spectral analysis assessment of pain control methods. *Pediatrics* 2009;**124**: 921–6. [DOI: 10.1542/peds.2009-0598]

Wisdorf-Houtkooper 1997 *{published data only}*
Wisdorf-Houtkooper KK. *The efficacy of pediatric pain management techniques for infants during inoculation procedures*. Michigan: Western Michigan University, 1997.

Zeltzer 1991 *{published data only}*
Zeltzer LK, Dolgin MJ, LeBaron S, LeBaron C. A randomized, controlled study of behavioral intervention for chemotherapy distress in children with cancer. *Pediatrics* 1991;**88**(1):34–42.

Additional references

Alexander 2003
Alexander J, Manno M. Underuse of analgesia in very young pediatric patients with isolated painful injuries. *Annals of Emergency Medicine* 2003;**41**(5):617–22.

Anand 2000
Anand KJ. Pain, plasticity and premature birth: a prescription for permanent suffering?. *Nature and Medicine* 2000;**6**:971–3.

Anand 2001
Anand KJS, International Evidence-Based Group for Neonatal Pain. Consensus statement for the prevention and management of pain in the newborn. *Archives of Pediatrics & Adolescent Medicine* 2001;**55**(2):173–80.

Baba 2010
Baba L, McGrath JM, Liu J. The efficacy of mechanical vibration analgesia for relief of heel stick pain in neonates: a novel approach. *Journal of Perinatal & Neonatal Nursing* 2010;**24**(3):274–83.

Barber 2004
Barber K. Pharmacology - noun. *The Canadian Oxford Dictionary - Oxford Reference Online*. University of Toronto Libraries: Oxford University Press, 2004.

Berde 2005
Berde CB, Jaksic T, Lynn AM, Maxwell LG, Soriano SG, Tibboel D. Anesthesia and analgesia during and after surgery in neonates. *Clinical Therapeutics* 2005;**27**(6): 900–21.

Bowlby 1969/1982
Bowlby J. *Attachment*. 2nd Edition. USA: Tavistock Institute of Human Relations, 1969/1982.

Brady-Fryer 2009
Brady-Fryer B, Wiebe N, Lander JA. Pain relief for neonatal circumcision. *Cochrane Database of Systematic Reviews* 2009, Issue 1. [DOI: 10.1002/14651858.CD004217.pub2]

Cepeda 2006
Cepeda MS, Carr DB, Lau J, Alvarez H. Music for pain relief. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD004843.pub2]

Cignacco 2007
Cignacco E, Hamers JPH, Stoffel L, van Lingen RA, Gessler P, McDougall J, et al.The efficacy of non-pharmacological interventions in the management of procedural pain in preterm and term neonates. A systematic literature review. *European Journal of Pain* 2007;**11**(2):139–52.

Cohen 1988
Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd Edition. Routledge, 1988.

Cohen 2008
Cohen LL, Lemanek K, Blount RL, Dahlquist LM, Lim CS, Palermo TM, et al.Evidence-based assessment of pediatric pain. *Journal of Pediatric Psychology* 2008;**33**(9):939–55.

Craig 2002
Craig KD, Korol CT, Pillai RR. Challenges of judging pain in vulnerable infants. *Clinics in Perinatology* 2002;**29**: 445–58.

Cyna 2010
Cyna AM, Middleton P. Caudal epidural block versus other methods of postoperative pain relief for circumcision in boys. *Cochrane Database of Systematic Reviews* 2010, Issue 11. [DOI: 10.1002/14651858.CD003005]

Deeks 2005
Deeks JJ, Higgins JPT, Altman DG, editors. Chapter 8: Analysing and presenting results. In: Higgins JPT, Green S editor(s). *Cochrane Handbook for Systematic Reviews of Interventions 4.2.5*. Chichester, UK: John Wiley & Sons, 2005.

Derbyshire 1999
Derbyshire SWG. Locating the beginnings of pain. *Bioethics* 1999;**13**(1):1–31. [MEDLINE: 11657057]

Field 1995
Field T. Infancy is not without pain. *Annals of Child Development* 1995;**10**:1–26.

Fitzgerald 2005
Fitzgerald M. The development of nociceptive circuits. *Nature Reviews Neuroscience* 2005;**6**:507–20.

Franck 2000
Franck L, Greenberg C, Stevens B. Pain assessment in infants and children. *Pediatric Clinics of North America* 2000;**47**(3):487–512.

Grunau 1987
Grunau R, Craig KD. Pain expression in neonates: facial action and cry. *Pain* 1987;**28**(3):395–410.

Grunau 1996
Grunau RVE, Whitfield M, Petrie JH. Children's judgments about pain at age 8-10 years: do extremely low birthweight children (<1000g) children differ from full birthweight peers?. *Journal of Child Psychology and Psychiatry* 1996;**39** (4):587–94.

Grunau 2000
Grunau RE. Long-term consequences of pain in human neonates. In: KJ Anand, BJ Stevens, PJ McGrath editor(s). *Pain in Neonates, 2nd Edition*. New York: Elsevier, 2000.

Higgins 2011
Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. www.cochrane-handbook.org.

Howard 2003
Howard RF. Current status of pain management in children. *JAMA* 2003;**290**(18):2464–9.

Johnston 1993
Johnston CC, Stevens B, Craig KD, Grunau RV. Developmental changes in pain expression in premature, full-term, two and four-month-old infants. *Pain* 1993;**52** (2):201–8.

Johnston 2011
Johnston CC, Fernandes AM, Campbell-Yeo M. Pain in neonates is different. *Pain* 2011;**152**(3):S65–S73.

Joint Committee 2000
Joint Fetus and Newborn Committee of the Candian Paediatric Society and American Academy of Pediatrics. Prevention and management of pain and stress in the neonate. *Pediatrics* 2000;**105**(2):454–61.

Kleiber 1999
Kleiber C, Harper D. Effects of distraction on children's pain and distress during medical procedures: a meta-analysis. *Nursing Research* 1999;**48**(1):44–9.

McClain 2005
McClainB, Kain ZN. Procedural pain in neonates: the new millennium. *Pediatrics* 2005;**115**(4):1073–5.

McGraw 1943
McGraw MD. *The neuromuscular maturation of the human infant*. New York: Columbia University Press, 1943.

Obeidat 2009
Obeidat H, Kahalaf I, Callister LC, Froelicher ES. Use of facilitated tucking for nonpharmacological pain management in preterm infants: a systematic review. *Journal of Perinatal and Neonatal Nursing* 2009;**23**(4):372.

Piira 2005
Piira T, Sugiura T, Champion GD, Donnelly N, Cole ASJ. The role of parental presence in the context of children's medical procedures: a systematic review. *Child: Care, Health and Development* 2005;**31**(2):233–43.

Pillai Riddell 2005
Pillai Riddell R, Stevens BJ, Katz J. Expanding the notion of chronic pain to infancy. Presentation at the Canadian Pain Society Annual Meeting. Pain Research & Management. 2005; Vol. 10:85.

Pillai Riddell 2006
Pillai Riddell RR, Stevens BJ. Non-pharmacological pain management in infancy. *The Encyclopedia of Pain*. Springer-Verlag, 2006.

Pillai Riddell 2009
Pillai Riddell R, Racine N. Assessing pain in infancy: the caregiver context. *Pain Research and Management* 2009;**14**: 27–32.

Prasopkittikun 2003
Prasopkittikun T, Tilokskulchai F. Management of pain from heel stick in neonates - an analysis of research conducted in thailand. *Journal of Perinatal and Neonatal Nursing* 2003;**17**(4):304–12.

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

Schechter 2007
Schechter NL, Zempsky WT, Cohen LL, McGrath PJ, McMurtry CM, Bright NS. Pain reduction during pediatric immunizations: Evidence-based review and recommendations. *Pediatrics* 2007;**119**(5):184–98.

Shah 2009
Shah PS, Aliwalas LL, Shah V. Breastfeeding or breast milk for procedural pain in neonates. *Cochrane Database of Systematic Reviews* 2009, Issue 1. [DOI: 10.1002/14651858.CD004950.pub2]

Simons 2003
Simons SHP, van Dijk M, Anand KS, Roofthooft D, van Lingen RA, Tibboel D. Do we still hurt newborn babies? : A prospective study of procedural pain and analgesia in neonates. *Archives of Pediatrics & Adolescent Medicine* 2003;**157**(11):1058–64.

Slifer 2002
Slifer KJ, Tucker CL, Dahlquist LM. Helping children and caregivers cope with repeated invasive procedures: how are we doing?. *Journal of Clinical Psychology in Medical Settings* 2002;**9**(2):131–52.

Stevens 1996
Stevens B, Johnston C, Petryshen P, Taddio A. Premature infant pain profile: development and initial validation. *Clinical Journal of Pain* 1996;**12**:13–22.

Stevens 2007
Stevens B, Pillai Riddell RR, Oberlander T, Gibbins S. Assessment of pain in neonates and infants. In: Anand KJ, Stevens BJ, McGrath PJ editor(s). *Pain in Neonates and Infants*. 3rd Edition. New York: Elsevier Limited, 2007.

Stevens 2010
Stevens B, Yamada J, Ohlsson A. Sucrose for analgesia in newborn infants undergoing painful procedures. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD001069.pub2]

Taddio 1997
Taddio A, Katz J, Ilersich AL, Koren G. Effect of neonatal circumcision on pain response during subsequent routine vaccination. *Lancet* 1997;**349**(9052):599–603.

Taddio 2010
Taddio A, Appleton M, Bortolussi R, Chambers C, Dubey V, Halperin S, et al.Reducing the pain of childhood vaccination: an evidence-based clinical practice guidelines (summary). *Canadian Medical Association Journal* 2010;**182**(18):1–7. [DOI: 10.1503/cmaj.092048]

Uman 2006
Uman LS, Chambers CT, McGrath PJ, Kisely SR. Psychological interventions for needle-related procedural pain and distress in children and adolescents. *Cochrane Database of Systematic Reviews* 2006, Issue 4. [DOI: 10.1002/14651858.CD005179.pub2]

Yates 2005
Yates SL, Morley S, Eccleston C, de C Williams AC. A scale for rating the quality of psychological trials for pain. *Pain* 2005;**117**:314–25.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

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

Akcan 2009

Methods	Study design: randomized between groups Study duration: 11 months
Participants	Total number: 50 Setting: NICU Diagnostic criteria: excluded if ventilated or sepsis Age: Minimum: 26 weeks Maximum: 28 weeks Mean: 31.6 weeks SD: 2.0 weeks Sex: Males: 54% Females: 46% Country: Turkey
Interventions	Total number of intervention groups: 1 Control group description: standard NICU care Total number of participants randomized to each of your groups: 25 Specific intervention: kangaroo care Intervention details: infants in intervention group were given kangaroo care for 45 minutes by their mothers for 45 minutes a day for 5 days. Infants wore a diaper and cap and placed between mother's breasts and covered with blanket. On the fifth day of the procedure, heel blood sample was taken at the 30th minute of kangaroo care and for 10 minutes after Integrity of intervention: satisfactory
Outcomes	Pain outcomes, time points when measured, and scale limits: **PIPP** 0 to 30 seconds Upper limit: 18 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Key conclusions of study authors: kangaroo care works for premies

Risk of bias

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria

Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial

Akcan 2009 (*Continued*)

Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	19 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Allen 1996

Methods	Study design: randomized between groups Total study duration: unknown
Participants	Total number: 285 Setting: hospital ambulatory pediatric clinic Diagnostic criteria: healthy neonates and infants Age: Minimum: 2 weeks Maximum: 18 months Mean: not reported SD: not reported Sex: Males: not mentioned Females: not mentioned Country: USA
Interventions	Total number of intervention groups: 1 Control group description: 1 Total number of participants randomized to each of your groups: 2 weeks old = 50 2 months old = 44 4 months old = 50 6 months old = 46

	9 months old = 28 15 months old = 30 18 months old = 37 Specific intervention: sterile water Intervention details: the infant orally received 2 mL of sterile water 2o minutes before needle Integrity of intervention: satisfactory
Outcomes	Pain outcomes, time points when measured, and scale limits: **Pain vocalization (cry)** Upper limit: 100% Lower limit: 0% High score =more pain (longer cry) Continuous
Notes	Adverse reactions: none Key conclusions of study authors: water administered prior to needle lowers pain response only when 1 needle is given is infants aged 2 weeks, 9 months and 18 months

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	High risk	0 - Insufficient information is reported to allow adequate comparisons to be made
Group equivalence	High risk	0 - Either equivalence of groups is not reported or there is evidence of non-equivalence
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person

Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	High risk	0 - The data analysis was not adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	High risk	7 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Axelin 2006

Methods	Study design: randomized cross-over trial Total study duration: 10 months (2003 to 2004)
Participants	Total number: 20 Setting: Turku University Hospital, NICU Diagnostic criteria: less than and equal to 37 weeks gestation, no major congenital anomalies, a need for regular endotracheal/pharyngeal suctioning, no analgesics for 4 hours before the procedure Age: Minimum: 24 weeks gestation, 6 days Maximum: 33 weeks gestation, 37 days Mean: not reported (median: 18 days) SD: not reported Sex: Males: 8 Females: 12 Country: Finland
Interventions	Total number of intervention groups: 1 Control group description: 1 Total number of participants randomized to each of your groups: 20 Specific intervention: facilitated tucking Intervention details: facilitated tucking by parents Integrity of intervention: satisfactory
Outcomes	Pain outcomes, time points when measured, and scale limits: **NIPS** Upper limit: 7 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: one infant got septicaemia after the experimental care and did not express pain during Funding sources: the South-Western Finnish Foundation of Neonatal Research Key conclusions of study authors: facilitated tucking by parents is an effective and safe pain management method during suctioning of preterm infants

Risk of bias

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met

Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported

Axelin 2006 (*Continued*)

Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	16 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Coders blinded to the order of conditions, but not to the condition itself
Selective outcome reporting	Unclear risk	Not specified
Other potential sources of bias	Low risk	There was no baseline inequality

Axelin 2009

Methods	Study design: randomized cross-over Study duration: 2 years
Participants	Total number: 22 Setting: NICU at university hospital Diagnostic criteria: preterm, but were excluded if they had unstable health conditions Age: Minimum: 23 weeks Maximum: 30 weeks Mean: 28 SD: 2.3 Sex: Males: 12 Females: 8 Country: Finland
Interventions	Total number of intervention groups: 1 Control group description: sterile: 2 mL of water on tongue Total number of participants randomized to each of your groups: 22 Specific intervention #1: facilitated tucking Intervention details: parent held infant in a side lying, flexed fetal type position offering support and skin contact (taught parent the procedure in advance) Integrity of intervention: Good
Outcomes	Pain outcomes, scale limits, and time points when measured: **PIPP** Measured for the first 30 seconds Higher scores = more pain

Axelin 2009 (*Continued*)

	Continuous	
Notes	Adverse reactions: none Key conclusions of study authors: facilitated tucking works	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change

Axelin 2009 (*Continued*)

Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	21 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Bauchner 1996

Methods	Study design: between groups Total study duration: unknown
Participants	Total number: 431 Setting: pediatric emergency department in a hospital Diagnostic criteria: infants aged 0 to 36 months coming to the pediatric ER, undergoing a venipuncture, intravenous cannulation, or urethral catheterization Age (page 863) Minimum: newborn Maximum: 3 years Mean: unknown SD: unknown Sex: Males: 246 (57%)

Bauchner 1996 (*Continued*)

	Females: 185 (43%) Country: USA Socio-demographics: 87% of the mothers were between the ages 20 and 24 years old; 22% were married; 43% had less than a high school degree Ethnicity: Black: 57.8%, 249 participants Hispanic: 21.3%, 92 participants White: 8.1%, 35 participants Other: 9.0%, 39 participants *does not add up to 431 participants. No explanation given as to why
Interventions	Total number of intervention groups: 2 Control group description: 1 (parents not present) Total number of participants randomized to each of your groups: 431 total (control: 131; intervention: 157; parent: 147) Specific intervention #1: intervention - parent instruction Intervention details: parents asked to sit at the head of the bed and talk to, touch, and maintain eye contact with their child Integrity of intervention: good Specific intervention #2: parent present Intervention details: parent present in the room, with no other instructions given Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: Cry (analyses) Measured during procedure (unspecified) in Hz More cry = more pain Continuous
Notes	Adverse reactions: none Funding sources: Maternal and Child Health Bureau Key conclusions of study authors: intervention (parent present, engaging in touch, talking and eye contact) during painful ER procedure are not effective in reducing pain (as measured by cry) in infants aged 0 to 36 months

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed

Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis

Bauchner 1996 (*Continued*)

Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Bellieni 2001

Methods	Study design: cross-over Total study duration: 3 months (December 2000 to February 2001)
Participants	Total number: 17 Setting: Siena Hospital NICU Diagnostic criteria: preterm infants Age: all infants were born less than 35 weeks, and the procedure took place within 10 days post natal life Minimum: 28 weeks Maximum: 35 weeks Mean: unknown SD: unknown Sex: Males: 7 Females: 10 Country: Italy
Interventions	Total number of intervention groups: 1 Control group description: no intervention whatsoever (just heel lance) Total number of participants randomized to each of your groups: 17 Specific intervention: water was administered orally 30 seconds before and during withdrawal Intervention details: water was instilled on the tip of the tongue with a syringe, introduced in mouth and moved to stimulate sucking. The amount that was administered was that necessary to maintain sucking (0.2 to 0.3 mL) until after the heel prick Integrity of intervention: satisfactory
Outcomes	Pain outcome, scale limits and time points when measured: **PIPP** Upper limit: 21 Lower limit: 0

	High score = more pain Continuous
Notes	Adverse reactions: none Key conclusions of study authors: non-nutritive sucking is effective in reducing pain for preterm infants undergoing heel lance when compared to receiving no treatment at all Miscellaneous comments by review authors: poor description of methods and time in which pain was measured; poor description of sample characteristics

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	High risk	0 - Insufficient information is reported to allow adequate comparisons to be made
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study

Bellieni 2001 (*Continued*)

Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	15 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Could see if baby received nothing or water
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Bellieni 2002

Methods	Study design: randomized between groups Total study duration: 5 months (February to June 2001)
Participants	Total number: 120 (60 participants meet the requirements of this review) Setting: hospital Diagnostic criteria: healthy newborns Age: Minimum: 38 weeks Maximum: 41 weeks Mean: not reported SD: 2.5

	Sex: Males: 31 Females: 29 Country: Italy Ethnicity: Caucasian
Interventions	Total number of intervention groups: 2 Control group description: no analgesic procedure Total number of participants randomized to each of your groups: 20 in each group (total of 60) Specific intervention #1: water during heel prick (sucking) Intervention details: the tip of 1 mL syringe without needle was placed in the baby's mouth and 1 mL of distilled water was given with gentle a movement of the syringe to stimulate sucking for 30 s before, during and after heelstick Integrity of intervention: good Specific intervention #2: sensorial saturation without glucose Intervention details: involved laying the infant on its side with legs and arms flexed but free to move; looking at the infant in the face, close up, to attract its attention and simultaneously massaging the infant's face and back; speaking to the infant softly but firmly; letting the infant smell the fragrance of the baby perfume on the physiotherapist's hands Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: Douleur Aiguë Nouveau-né **(DAN) scale** 30 seconds pre heelstick 30 seconds post heelstick Upper limit: 10 Lower limit: 0 Higher score = more pain Continuous
Notes	Adverse reactions: none Key conclusions of study authors: sensorial stimulation without glucose does not produce an analgesic effect in healthy newborns during the 30 second post heel prick but rather, it increased irritation and awareness Sucking (water oral) does significantly reduce pain scores during 30 seconds post heel lance in healthy newborn infants

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met

Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	High risk	0 - Insufficient information is reported to allow adequate comparisons to be made
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis

Bellieni 2002 (*Continued*)

Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	20 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Blass 1999

Methods	Study design: randomized between groups Total study duration: 12 months (June 1997 to June 1998)
Participants	Total number: 40 (20 meet the requirements of this review) Setting: Hospital (Boston Medical) Diagnostic criteria: healthy newborns Age: Minimum: 34 hours old Maximum: 55 hours old Mean: not reported SD: not reported Sex: Males: 17 Females: 23 Country: United States Ethnicity: 55% African American; 22.5% White; 12.5% Hispanic; 2.5% Asian; 2.5% American Indian; 5% unclassified
Interventions	Total number of intervention groups: 1 - pacifier dipped in water Control group description: water only (delivered by syringe) Total number of participants randomized to each of your groups: 20 total (10 in each group) Specific intervention #1: infant was given either water only or a pacifier prior to heel lance Intervention details: 4 minutes pre heel lance, the infant was given a heat pad to warm feet and was wrapped in a blanket and turned supine to his or her bassinet. 60 seconds before the heelstick, the infant was given either water only (delivered by syringe) or pacifier (coated in water). Pacifier was redipped in water every 30 seconds. The heat pad was then removed and the heel lance occurred lasting 1 to 3 minutes. A band-aid was then placed on the wound Integrity of intervention: good

Outcomes	Pain outcomes, scale limits, and time points when measured: **Crying** Heel lance (2 minutes post) Recovery (5 minutes post) Upper limit: 100% Lower limit: 0% High score = more pain Continuous **Grimacing (furrowing of brow and eye squinting)** Heel lance (2 minutes post) Recovery (2 minutes post) Upper limit: 100% Lower limit: 0% Continuous
Notes	Adverse reactions: none Funding sources: National Institute of Mental Health Key conclusions of study authors: sucking a water dipped pacifier caused modest and variable reductions in grimacing during the 2 minute post heel lance in healthy neonates. Further analyses revealed that this intervention is only effective if the sucking rate is greater than 30 times/minute prior to the heel lance Miscellaneous comments by review authors: results are not significant and no data are given at the recovery period (2 mins to 5 mins post heel lance)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history

Group equivalence	High risk	0 - Either equivalence of groups is not reported or there is evidence of non-equivalence
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	14 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Not blind to pacifier versus syringe condition

Blass 1999 (*Continued*)

Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	Low risk	Some participants had been circumcised the day before the study Some had heel lance before These participants were randomly distributed to both groups

Bo 2000

Methods	Study design: cross-over design Total study duration:
Participants	Total number: 27 Setting: special care baby unit, NICU Diagnostic criteria: healthy neonates and premature neonates Age: Minimum: 30 weeks Maximum: 41 weeks Mean: 37 weeks SD: 3.43 weeks Sex: Males: 17 Females: 10 Country: Hong Kong/China Ethnicity: Chinese
Interventions	Total number of intervention groups: 1 Control group description: infants were placed in the supine position. Total number of participants randomized to each of your groups: 27 Specific intervention: non-nutritive sucking Intervention details: infants were given a small, standard, short and hollow soft latex nipple. Gentle pressure was applied to keep it in the infant's mouth Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIPS** 1 minute pre heelstick 13 minutes post heelstick Upper limit: 7 Lower limit: 0 Higher score = higher pain Continuous
Notes	Adverse reactions: none Key conclusions of study authors: non-nutritive sucking is an effective intervention for pain regulation after heelstick in neonates

Bo 2000 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	High risk	0 - Either equivalence of groups is not reported or there is evidence of non-equivalence
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori

Bo 2000 (*Continued*)

Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Bustos 2008

Methods	Study design: randomized between groups Total study duration: unknown
Participants	Total number: 62 Setting: Sydney Children's Hospital Outpatient Immunization Clinic Diagnostic criteria: adequate grasp of English, spoke English to infant Age: Minimum: 5 months Maximum: 7 months Mean: not reported SD: not reported Sex: Males: 24 (in the final sample) Females: 26 (in the final sample) Country: Australia

Interventions	Total number of intervention groups: 1 Control group description: standard care Total number of participants randomized to each of your groups: intervention = 30; control group = 32 (total = 62) Specific intervention: coping - promoting verbalizations Intervention details: intervention group received a one-page information sheet with information about 3 types of parental verbalizations associated with lower pain outcomes for infants. Parents were encouraged to review the sheet before their appointment. All infants were placed in supine position on the treatment table. The injection site was cleaned with an alcohol swab and the injection was then performed. During the injection, parents held infant's leg Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NFCS** Injection phase: 15 seconds post needle Recovery phase: 15 seconds following injection phase (scored in 5 second blocks) Upper limit: 30 Lower limit: 0 High score = higher expression of pain Continuous
Notes	Adverse reactions: none Funding sources: Cavenadian Institutes of Health Research (Pain In Child Health Strategic Training Initiative); Mayday Foundation Key conclusions of study authors: the promotion of parent-coping verbalizations during an infant's 6 month immunization is able to reduce the duration of infant cry following an injection There was no significant effect of intervention on NFCS outcome measure

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant

Sample characteristics	High risk	0 - Insufficient information is reported to allow adequate comparisons to be made
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	Low risk	18 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)

Bustos 2008 (*Continued*)

Blinding	Low risk	All coders were blinded to group assignment
Selective outcome reporting	Low risk	Cry duration
Other potential sources of bias	Low risk	Did not examine possible differences between ethnicity or SES

Campos 1994

Methods	Study design: between groups Total study duration: unknown
Participants	Total number: 60 Setting: mid-western neo-natal nursery Diagnostic criteria: full-term birth; Apgar scores > 7; absence of respiratory or other health problems Age: Minimum: not reported Maximum: not reported Mean: 51.5 hrs SD: 12 hrs Sex: Males: 30 Females: 30 Country: United States
Interventions	Total number of intervention groups: 2 Control group description: infants received no comforting care after heelstick; were placed prone in bassinet after heelstick and covered with a blanket Total number of participants randomized to each of your groups: 20 participants in each of the 3 groups, total number of participants is 60 Specific intervention #1: rocking Intervention details: infants were held at the experimenter's shoulder, rocked in a rocking chair at a rate of 30 cycles per minute (auditory signal fed to experimenter) post procedure Integrity of intervention: good Specific intervention #2: pacifier Intervention details: infants were held in the experimenter's lap while the pacifier was administered for 8 minutes after heelstick; gentle pressure was applied to keep that pacifier in the infant's mouth Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **Latencies to cry cessation** 0 to 2 minutes post heelstick 3 to 4 minutes post heelstick Upper limit: 180 seconds Lower limit: 0 seconds Higher score = longer latency to cry cessation

Campos 1994 (*Continued*)

	Continuous	
Notes	Adverse reactions: none Funding sources: National Association of Neonatal Nurses; The University of Illinois Key conclusions of study authors: newborns benefit from both rocking and pacifiers when undergoing heelstick procedures	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid

Campos 1994 (*Continued*)

Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	High risk	13 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	Low risk	Measured cry from audiotapes, so coders were unaware of group assignment
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Carbajal 1999

Methods	Study design: between groups Total study duration: April 1997 to June 1997
Participants	Total number: 150 Setting: maternity ward of a hospital in France Diagnostic criteria: newborns who had a 5-minute Apgar score greater than or equal to 7, were medically stable, had not received naloxone during the previous 24 hours, and were not fed in the previous 30 minutes Age: Minimum: 37 weeks

Carbajal 1999 (*Continued*)

	Maximum: 42 weeks Mean: 40 weeks SD: not reported Sex: Males: 89 Females: 61 Country: France
Interventions	Total number of intervention groups: 5 (but only 1 does not use sucrose) Control group description: placebo, 2 mL of sterile water Total number of participants randomized to each of your groups: 100 (does not indicate how many were in each group) Specific intervention #1: pacifier Intervention details: pacifier given to baby Integrity of intervention: unknown
Outcomes	Pain outcomes, scale limits, and time points when measured: **DAN** (Douleur Aiguë Nouveau-né) Article does not indicate what time points were measured Upper limit: 10 Lower limit: 0 Higher score = more pain Continuous
Notes	Adverse reactions: none Funding sources: no external funding Key conclusions of study authors: pacifiers were effective in reducing infant pain

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	High risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history

Group equivalence	High risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient e.g. outcomes collected by therapist
Justification of outcomes	High risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	High risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	High risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	High risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	High risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	22 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)

Carbajal 1999 (*Continued*)

Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	Unclear risk	Very short article, not a lot of available information

Carbajal 2003

Methods	Study design: between groups Total study duration: 5 months
Participants	Total number: 180 Setting: maternity ward of hospital Diagnostic criteria: equal to or greater than 37 weeks gestation; had APGAR scores of 7 or higher at 5 minutes; were greater than or equal to 24 hours old; undergoing venipuncture; breast-fed; had not been fed for the previous 30 minutes; no medical instability; no medication in the previous 24 hours Age: Minimum: 39.6 weeks Maximum: 40 weeks Mean: 39.8 SD: 1.23 Sex: Males: 93 Females: 87 Country: France
Interventions	Total number of intervention groups: 3 (we will only focus on the first one for this review; other interventions included breast milk and sucrose) Control group description: Infants given sterile water as a placebo Total number of participants randomized to each of your groups: 45 in each group (180 total) Specific intervention: held in mother's arms without breast feeding Intervention details: held in mother's arms, without breast feeding Integrity of intervention: satisfactory
Outcomes	Pain outcome, scale limits and time points when measured: **Douleur (DAN) scale** Time points: unknown Upper limit: 10 Lower limit: 0 High score = more pain Continuous **PIPP** Time points: unknown Upper limit: 21 Lower limit: 0

	Higher score = more pain Continuous	
Notes	Adverse reactions: none Funding sources: Foundation Pour La Sante, France Key conclusions of study authors: there were no significant differences in pain reactivity between newborns being held in mother's arms or newborns who were given sterile water	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified

Carbajal 2003 (*Continued*)

Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	Low risk	20 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Coders knew group assignment
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Castral 2008

Methods	Study design: between groups Total study duration: 8 months
Participants	Total number: 59 Setting: interediary neonatal care unit Diagnostic criteria: preterm neonate Age: Mean: 251.35 days Sex: Males: 31

	Females: 28 Country: Brazil
Interventions	Total number of intervention groups: 1 Control group description: Infant was placed in a lateral decubitus poistion with their heads raised by a cloth diaper in their cribs or incubators, wearing only diapers and rolled-up in blankets Total number of participants randomized to each of your groups: 28 to control, 31 to skin-to-skin. Specific intervention: skin-to-skin contact Intervention details: The infants were held by their mothers in an eleated prone position, at approxcimately 60 degrees across mothers' chest while mothers' hands remained around the babies' backs, hidden under blankets Integrity of intervention: satisfactory
Outcomes	Pain outcome, scale limits and time points when measured: **Neonatal Facial Coding Systems (NFCS)** Time points: baseline (120 seconds), treatment (120 seconds), heel cleaning (20 seconds), heel prick (20 seconds), wound compression (20 seconds), recovery (120 seconds) Upper limit: 10 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: none mentioned Key conclusions of study authors: skin-to-skin contact promoted reduction in behavioural measures at heel prick but not at recovery

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias

Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	High risk	0 - Randomization is mentioned but there is not an adequate description of the methods used
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	0 - Power calculations were not conducted a priori.
Sufficient sample	Low risk	0-A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness

Castral 2008 (*Continued*)

Total	Unclear risk	15

Catelin 2005

Methods	Study design: cross-over Total study duration:
Participants	Total number: 45 Setting: NIDCAP-reliable NICU Diagnostic criteria: participants were excluded if: they received treatment with a muscle relaxant, sedative, antiepileptic, or analgesic drug (except sucrose) during the last 24 hours, a congenital defect, a neurological abnormality including convulsion, intraventricular hemorrhage grade higher than II according to the Papile scale, and periventricular leukomalacia Age: Minimum: less than or equal to 32 weeks Maximum: greater than or equal to 37 weeks Mean: 34.47 weeks SD: 1.0 weeks Sex: Males: 27 Females: 18 Country: France
Interventions	Total number of intervention groups: 1 Control group description: no EBI, no specific protection from light or noise, in the supine position and without swaddling or any postural support Total number of participants randomized to each of your groups: 45 Specific intervention: Environmental and Behavioral Interventions (EBI) Intervention details: EBI included attenuated noise and light with closed doors and covered incubator, lateral posture with head, back, and feet contacting supportive bedding, and opportunity for grasping or sucking Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIPS** 2 minutes before weighing, during weighing and 5 and 30 minutes after weighing Upper limit: 7 Lower limit: 0 High score = more pain Continuous **EDIN** 2 minutes before and 5 and 30 minutes after weighing Upper limit: 15 Lower limit: 0 High score = more pain Continuous

Notes	Adverse reactions: none Funding sources: grants from the French ministry of health, the Foundation de France, SESEP and the Fondation CNP Key conclusions of study authors: EBI procedures were associated with lower heart rates just after weighing procedures

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid

Catelin 2005 (*Continued*)

Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori.
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	16 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Chermont 2009

Methods	Study design: randomized between subject Total study duration: March 2006 to October 2007
Participants	Total number: 640 Setting: hospital Diagnostic criteria: healthy term infants with no clinical problems Age: Minimum: 37 weeks Maximum: 41 6/7 weeks Mean: 39 weeks SD: 1 week

	Sex: Males: 320 Females: 320 Country: Brazil
Interventions	Total number of intervention groups: 1 Control group description: 1 mL of sterile water was given on tongue 2 minutes before cleansing Total number of participants randomized to each of your groups: 160 Specific intervention: skin to skin contact Intervention details: infant wearing only a diaper was placed on bare chest of mother and held for 2 minutes. 1 mL of sterile water was administered to tongue Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NFCS** Reactivity and immediate regulation Upper limit: 10 Lower limit: 0 Continuous
Notes	Adverse reactions: none Funding sources: no funding sources relevant to this study Key conclusions of study authors: kangaroo care is no different than a no treatment control for pain reactivity or pain regulation for neonates

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history

Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	20 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified

Chermont 2009 (*Continued*)

Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Cohen 2002

Methods	Study design: between groups Total study duration: 1 year
Participants	Total number: 90 Setting: rural clinic Diagnostic criteria: healthy babies Age: Minimum: 2 months Maximum: 3 years Mean: 12 months SD: 8.6 months Sex: Males: 44 Females: 46 Country: United States of America Co-morbidity: none Socio-demographics: income average was $24,000/year Parents: 2 years of college Ethnicity: 88% were Caucasian
Interventions	Total number of intervention groups: 2 Control group description: typical care, interact normally with infant but without movie or toy distraction Total number of participants randomized to each of your groups: Specific intervention: video distraction Intervention details: Teletubbies movie and/or toys Quality of intervention: poor
Outcomes	Pain outcomes, scale limits, and time points when measured: **MBPS** Injection and recovery phase Upper limit: 40 Lower limit: 0 High score = more pain Continuous **VAS - Parent** During and 3 minutes after immunization Upper limit: 100 Lower limit: 0 High score = more pain Continuous

Cohen 2002 (*Continued*)

	VAS - Nurse During and 3 minutes after immunization Upper limit: 100 Lower limit: 0 Higher score = more pain Continuous
Notes	Adverse reactions: none Funding sources: none mentioned Key conclusions of study authors: nurse-directed distraction works but not sure if it is the toy, nurse, parent or combination

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	High risk	0 - The inclusion/exclusion criteria are not clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	High risk	0 - There is no clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	High risk	0 - Randomization is mentioned but there is not an adequate description of the methods used
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist

Cohen 2002 (*Continued*)

Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	High risk	10 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	Low risk	Nurses, parents, and coder blind to hypotheses
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	Low risk	Major missing data unaccounted for

Cohen 2006

Methods	Study design: between groups Total study duration: 5 months
Participants	Total number: 136 Setting: university affiliated medical centre; private practice office Diagnostic criteria: any English-speaking families with an infant between the ages of 1 and

Cohen 2006 (*Continued*)

	24 months present for routine checkups and vaccinations Age: Minimum: 1 month Maximum: 21 months Mean: 7.6 months SD: 5 months Sex: Males: 55 Females: 80 Country: United States Socio-demographics: income ranged from $0 to $170,000, M = $38,740, SD = $29,158 Parent's education ranged from 8th grade-post baccalaureate, most had completed 1 year of college Ethnicity: 90.4% Caucasian; 2.2% Native American; 1.5% African American; 0.7% Hispanic; 5.2% other
Interventions	Total number of intervention groups: 1 Control group description: parent and nurse were encouraged to interact with the infant in their usual manner, no movie or toy distraction provided Total number of participants randomized to each of your groups: approximately 63 per group (total = 136) Specific intervention: distraction - video Intervention details: prior to data collection, nurses engaged in brief intervention training and parents were briefly instructed in distraction techniques. During the immunization, a DVD movie (choice between 2) played on a hand-held DVD player 6 inches from the child. Parents were instructed that they could redirect attention to DVD Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **MAISD (infant distress, crying, screaming or flailing)** Coded for 5-second periods up to 2 minutes post needle Upper limit: 1 Lower limit: 0 High score = higher pain Continuous
Notes	Adverse reactions: none Funding sources: Mayday Fund Key conclusions of study authors: video distraction is a simple and practical means of providing distress relief to infants during routine injections

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria

Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial

Cohen 2006 (*Continued*)

Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Coders could see (in the video) which group was used for coding
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Comaru 2009

Methods	Study design: cross-over Total study duration: April to October 2004
Participants	Total number: 47 Setting: NICU Diagnostic criteria: preterm Age: Minimum: not reported Maximum: 35 weeks Mean: 32 weeks SD: 2 weeks Sex: Males: not reported Females: not reported Country: Brazil Co-morbidity: respiratory distress syndrome was diagnosed in approximately 35% of the studied babies Socio-demographics: not reported Ethnicity: not reported
Interventions	Total number of intervention groups: 1 Control group description: babies' diaper changed in the incubator Total number of participants randomized to each of your groups: 47 Specific intervention: nesting

	Intervention details: slightly flexed posture, side lying, limbs directed to midline: head, back, links feet with folded up towel Integrity of intervention: not needed
Outcomes	Pain outcomes, scale limits, and time points when measured: **NFCS** Upper limit: 1 Lower limit: 0 Continuous
Notes	Adverse reactions: none Funding sources: Research Support Fund (FIPE) of the Hospital of Clinicas in Porto Alegre and by the CNPq Scholarship Key conclusions of study authors: nesting works in diminishing distress during diaper changes

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants

Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	21 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Corbo 2000

Methods	Study design: cross-over design Total study duration: unknown
Participants	Total number: 26 Setting: NICU of the Division of Neonatology (University of Naples) Diagnostic criteria: neonates without severe complications Age: Minimum: 2 days Maximum: 15 days Mean: 4.9 days SD: 3.3 days Sex: Males: 16 Females: 10 Country: Italy
Interventions	Total number of intervention groups: 1 Control group description: infant in the supine position with their head towards the examiner, in incubator/crib/room, no pacifier Total number of participants randomized to each of your groups: 26 Specific intervention: non-nutritive sucking (pacifier) Intervention details: the pacifier was placed in the infant's mouth 3 minutes pre procedure and for 3 minutes post procedure Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **Brazelton State 6** 3 minutes post procedure Upper limit: 180 seconds Lower limit: 0 seconds Higher score = longer time spent in distressed state Continuous
Notes	Adverse reactions: none Funding sources: none reported Key conclusions of study authors: non-nutritive sucking can reduce the period of behavioral distress in newborn, premature infants following heelstick blood sampling

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met

Corbo 2000 (*Continued*)

Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported

Corbo 2000 (*Continued*)

Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	15 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Could see pacifier while coding
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Corff 1995

Methods	Study design: cross-over Study duration: unknown
Participants	Total number: 36 Setting: NICU Diagnostic criteria: less than 22 days old, and no anomalies Age: Minimum: 25 weeks Maximum: 35 weeks Mean: 30 weeks SD: unknown Sex: Males: not reported Females: not reported Country: USA
Interventions	Total number of intervention groups: 1 Control group description: normal nursery routine Total number of participants randomized to each of your groups: 30 Specific intervention: facilitated tucking Intervention details: facilitated tucking was supplied during and after the heelstick by one consistent neonatal nurse Integrity of intervention: poor (not reported)
Outcomes	Pain outcomes, scale limits, and time points when measured: **Total crying time** Upper limit: 100%

	Lower limit: 0 % Longer cry = more distress Continuous	
Notes	Adverse reactions: none Funding sources: not reported Key conclusions of study authors: facilitated tucking works	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	High risk	0 - There is no clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	High risk	0 - Randomization is mentioned but there is not an adequate description of the methods used
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified

Corff 1995 (*Continued*)

Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	14 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Cramer-Berness 2005

Methods	Study design: between groups Total study duration: 12 months
Participants	Total number: 123 Setting: health care clinic in upstate New York Diagnostic criteria: healthy infants between 2 months and 24 months of age who were receiving routine immunizations Age:

	Minimum: 2 months Maximum: 24 months Mean: 8.70 months SD: 5.92 months Sex: Males: 60 Females: 63 Country: United States Co-morbidity: none Socio-demographics: never married = 59%; married = 28.7%; separated = 9%; divorced = 3.3% Less than $10,000 = 32.5%; $10,000 to $20,000 = 29.8%; $20,001 to $30,000 = 24.6%; greater than $30,001 = 13.2%; did not answer = 7.3% Ethnicity: Caucasian parents = 63.9% Caucasian infants = 51.3%; Hispanic parents = 13.4% Hispanic infants = 7.6%; African American parents = 12.6% African American infants = 15.1%; other parents = 10.1% other infants = 8.4%
Interventions	Total number of intervention groups: 2 Control group description: typical care - parents did not receive prompting to engage in usual behaviors or distraction Total number of participants randomized to each of your groups: typical care = 41, supportive care = 42, distraction = 40 (total = 123) Specific intervention #1: supportive care Intervention details: parents were asked what they normally do to reduce infant distress. They were encouraged by the researcher to engage in these techniques Integrity of intervention: satisfactory Specific intervention #2: distraction Intervention details: parents were encouraged to use a distraction toy and to direct infant's attention towards the toy (e.g. "Look at this!") Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **MBPS** 5 seconds pre-needle 5 seconds during needle 5 seconds after needle Upper limit: 3 Lower limit: 0 High score = more distress Continuous
Notes	Adverse reactions: 3 parents withdrew from the study (no reason given) Funding sources: not reported Key conclusions of study authors: brief prompting for parents to use their normal coping strategies may be more effective than training in helping infants recover more quickly from routine procedure pain
Risk of bias	

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori

Cramer-Berness 2005 (*Continued*)

Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	18 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Could see toy on video
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Cramer-Berness 2005b

Methods	Study design: randomized between groups Total study duration: 14 months
Participants	Total number: 117 Setting: pediatrician's office Diagnostic criteria: healthy infants Age: Minimum: 2 months Maximum: 24 months Mean: 8.61 months SD: 5.78 months Sex: Males: 63 Females: 54 Country: United States Socio-demographics: Annual household income:

	Less than $10,000: 29.1% $10,000 to $20,000: 33.3% $20,001 to $30,000: 10.3% $30,001 to $50,000: 10.3% $50,001 and higher: 6.8% Ethnicity: Caucasian parents: 64.1%; Caucasian infants: 55.6% Hispanic parents: 8.5%; Hispanic infants: 6.0% African American parents: 14.5%; African American infants: 12.0% Asian parents: 1.7%; Asian infants: 1.7% Native American parents: 2.6%; Native American infants: 1.7% Multiracial parents: 7.7%; Multiracial infants: 23.1%
Interventions	Total number of intervention groups: 2 Control group description: typical care Total number of participants randomized to each of your groups: 117 Specific intervention #1: audiovisual distraction Intervention details: parents were instructed to encourage infants to focus on toys throughout immunization process by verbal statements such as "look at this!" and engaging behaviors, such as playing music with the toy Integrity of intervention: good Specific intervention #2: tactile distraction Intervention details: parents were instructed to tickle their infants during the immunization and alternate where they tickled with each immunization Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **MBPS** 5 seconds post needle Upper limit: 10 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: none reported Key conclusions of study authors: distraction (auditory or tactile) does not reduce pain in healthy infants aged 2 to 24 months during immunization Miscellaneous comments by review authors:if more than 1 needle was given they averaged out the pain score between needles to account for 1 to 4 needles. Therefore, there was a separate score for "during needle" phase

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	High risk	0 - The inclusion/exclusion criteria are not clearly specified for the sample and there is evidence of adherence to the criteria

Cramer-Berness 2005b (*Continued*)

Evidence that the criteria has been met	High risk	0 - There is no clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial

Cramer-Berness 2005b (*Continued*)

Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	Low risk	19 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Curtis 2007

Methods	Study design: between groups; comparison to baseline Total study duration: 6 months
Participants	Total number: 84 Setting: pediatric Emergency Department at the Stollery Children's Hospital in Edmonton, Alberta, Canada Diagnostic criteria: all infants up to 6 months corrected age that required venipuncture; participants required to have had nothing by mouth for 5 minutes prior to venipuncture; could not be critically ill; fructose intolerance; no EMLA at the site of venipuncture Age: Minimum: 0 months Maximum: 6 months Mean: 58.25 days SD: 56.50 days Sex: Males: 44 Females: 40 Socio-demographics: not reported Ethnicity: not reported
Interventions	Total number of intervention groups: 1 (placebo, and placebo and pacifier) Control group description: placebo group received 2 mL of sterile water, administered via syringe over tongue at 2 minutes prior to venipuncture Total number of participants randomized to each of your groups: placebo = 19, pacifier and placebo = 22

	Specific intervention: timers used to co-ordinate events; solution administered by research nurse to anterior aspect of tongue over 30 seconds via syringe and pacifier inserted orally (if deemed necessary) at 2 minutes post intervention. Venipuncture was performed by nurse. Parents interacted with voice and touch as per usual Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **FLACC Pain Scale** Measured before and after procedure (30 to 60 seconds post intervention) Change from baseline = outcome Upper limit: 10 Lower limit: 0 High score = higher pain Continuous
Notes	Adverse reactions: none Funding sources: pediatric residents training committee research award at the Department of Pediatrics, Stollery Children's Hospital Key conclusions of study authors: pacifiers are effective analgesics Miscellaneous comments from the study authors: "Parents interacted with voice or touch as per normal" (page 4). This may have confounded results

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data

Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	21 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting

Other potential sources of bias	High risk	There were no other potential sources of bias reported

de Sousa 2008

Methods	Study design: between groups Total study duration: unknown
Participants	Total number: 105 Setting: Maranhao Federal University Hospital NICU Diagnostic criteria: excluded neonates receiving invasive or non-invasive ventilation, a chest drain, oxygen therapy or a tracheotomy, hemodynamically unstable neonates, neonates who had received analgesia or sedative medications during the 48h period prior to the study, neonates with congenital anomalies, chromosomal syndrome, ventricular hemorrhage or periventricular leukomalacia, and neonates of mothers taking illicit drugs Age: Minimum: 28 weeks gestational age Maximum: 36 weeks gestational age Mean: 4.33 days SD: 2.54 days Sex: Males: not reported Females: not reported Country: Brazil
Interventions	Total number of intervention groups: 2 (though we will only focus on the kangaroo intervention, as the second intervention focuses on glucose) Control group description: neonates placed in the prone position at a 30 to 45 degree angle of elevation for 10 minutes pre heel lance procedure Total number of participants randomized to each of your groups: 105 (control = 35, kangaroo = 35, glucose = 35) Specific intervention: kangaroo care Intervention details: infant was placed with skin-to-skin contact on their mother's chest 10 minutes before heel lance procedure The mother was sitting comfortably at a 45 to 60 degree angle Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **Brow bulge** 30 seconds post heel lance Upper limit: not reported Lower limit: not reported Higher score = more pain Continuous **Eye squeeze** 30 seconds post heel lance Upper limit: not reported Lower limit: not reported

	Higher score = more pain Continuous **Nasolabial furrow** 30 seconds post heel lance Upper limit: not reported Lower limit: not reported Higher score = more pain Continuous
Notes	Adverse reactions: none Funding sources: none reported Key conclusions of study authors: skin-to-skin contact is a low-cost method that can be applied to almost all the neonates hospitalized in the NICU

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	High risk	0 - Insufficient information is reported to allow adequate comparisons to be made
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported

de Sousa 2008 (*Continued*)

Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	19 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Elserafy 2009

Methods	Study design: cross-over Study duration: January 2005 and May 2007

Elserafy 2009 (*Continued*)

Participants	Total number: 36 Setting: NICU tertiary care Diagnostic criteria: exclusion criteria were 1) exposure to maternal sedation; 2) occurrence of any procedure performed within 24 hours in preterm infants whose mothers had had general anesthesia during delivery; 3) the presence of major neurologic abnormalities; 4) Apgar scores at 5 minutes of greater than 5; 5) presence of necrotizing intestinal colitis; 6) nothing by mouth status for any reason; 7) being preterm with hyperglycemia Age: Minimum: 27 weeks Maximum: 36 weeks Mean: 32.4 weeks SD: 2.9 Sex: Males: not reported Females: not reported Country: Saudi Arabia
Interventions	Total number of intervention groups: 3 Control group description: standardized care Total number of participants randomized to each group: 36 Specific intervention: water pacifier Intervention details: infants were given pacifiers coated in 0.5 mL of water 2 minutes prior Integrity of intervention: good Specific intervention #2: pacifier alone Intervention details: standard nipple stuffed with gauze square for resistance, held in infant's mouth for 2 minutes prior Integrity of intervention: good Specific intervention #3: sterile water Intervention details: 0.5 mL of sterile water without pacifier Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **PIPP** 15 seconds after the venipuncture Upper limit: 3 Lower limit: 0 Higher score = more pain Continuous
Notes	Adverse reactions: none Funding sources: none reported Key conclusions of study authors: water pacifier, pacifier alone, and sterile water were not as effective as sucrose and pacifier

Risk of bias

Bias	Authors' judgement	Support for judgement

Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained

Elserafy 2009 (*Continued*)

Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	18 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Fearon 1997

Methods	Study design: cross-over Total study duration: not reported
Participants	Total number: 15 Setting: NICU Diagnostic criteria: Age: Minimum: 27 weeks Maximum: 36 weeks Mean: younger group (29.6 weeks) and older group (33.7 weeks) SD: younger group (0,90 weeks) and older group (2.16 weeks) Sex: Males: 9 Females: 6 Country: Canada Socio-demographics: not collected Ethnicity: not collected
Interventions	Total number of intervention groups: 1 Control group description: 1

Fearon 1997 (*Continued*)

	Total number of participants randomized to each group: 15 Specific intervention: swaddling Intervention details: placed infant in a supine position on a soft clean cloth. Infant's arms were crossed over in a relaxed position and tucked in on sides Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NFCS** Pre needle, post needle and after needle Upper limit: 10 Lower limit: 0 Higher score = more pain
Notes	Adverse reactions: none Funding sources: Ontario Ministry of Health Nursing Innovation Fund grant and the Ontario Ministry of Health Career Scientist Award Key conclusions of study authors: pain behaviors reduced by swaddling

Risk of bias

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person

Fearon 1997 (*Continued*)

Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori.
Sufficient sample	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Planned analysis	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	20 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Ferber 2008

Methods	Study design: cross-over design, within subject Total study duration: unknown
Participants	Total number: 31 Setting: NICU Diagnostic criteria: healthy mothers with single pregnancies and documented prenatal care, who were admitted at 28 to 34 weeks of gestation with premature uterine contractions and delivered vaginally Age: Minimum: 5 days Maximum: 10 days Mean: not reported SD: not reported Sex: Males: not reported Females: not reported Country: Israel
Interventions	Total number of intervention groups: 1 Control group description: blood test and in standard crib care Total number of participants randomized to each group: 30 Specific intervention: kangaroo care Intervention details: infants were undressed and placed on the mother's chest with skin-to-skin contact. The mother and infant were covered not tightly by blankets and infants were positioned in a flexed position. Mothers were bedded in a comfortable position in an armchair in a silent and semi-dark corner of the NICU Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIDCAP** During test (2 minutes) and post test (10 minutes) Upper limit: 10 (in minutes) Lower limit: 0 Higher score = longer duration Continuous
Notes	Adverse reactions: none Funding sources: none reported Key conclusions of study authors: kangaroo care lessens painful reactions after blood test procedures in premature infants

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria

Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	High risk	0 - Either equivalence of groups is not reported or there is evidence of non-equivalence
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial

Ferber 2008 (*Continued*)

Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	19 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	Low risk	Does not report differences in group equivalence

Gormally 2001

Methods	Study design: randomized between groups Total study duration: unknown
Participants	Total number: 94 Setting: community hospital Diagnostic criteria: healthy infants born at term Age: Minimum: 37 weeks Maximum: not reported Mean: 39.3 SD: 1.3 Sex: Males: 38 Females: 47 Country: Canada Socio-demographics: mothers in control group had a mean age of 29.9 years and 10.7 years of education; mothers in holding group had a mean age of 29.4 years and 11.4 years of education
Interventions	Total number of intervention groups: 1 Control group description: no holding and water taste Total number of participants randomized to each of your groups: 94 Specific intervention: holding and water taste Intervention details: the infant is held in the arms of a female RA beginning 4 minutes pre

	heelstick, given sterile water and rested in the morning Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NFCS (% of the time)** 1 minute post heel lance, 2 minutes post heel lance Upper limit: 100 Lower limit: 0 Higher score = more pain Continuous **Cry (% of the time)** 1, 2, and 3 minutes post heel lance Upper limit: 100 Lower limit: 0 Higher score = more time spent crying/more pain Continuous
Notes	Adverse reactions: none Funding sources: Medical Research Council of Canada, Allan Ross Fellowship, Lewis Sessenwein Academic Award Key conclusions of study authors: providing care-giving context reduces pain in neonates

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data

Gormally 2001 (*Continued*)

Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	Low risk	14 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Gray 2000

Methods	Study design: between groups Total study duration: 8 months
Participants	Total number: 30 Setting: Boston Medical Centre in Boston, Massachusetts Diagnostic criteria: infants 33 to 55 hours old (37 weeks plus gestation); healthy (no congenital abnormalities, medical complications or drug exposure); first heelstick procedure for all infants; no infants required oxygen or ventilatory support Age: Minimum: 33 hours Maximum: 55 hours Mean: not reported SD: not reported Sex: Males: 11 Females: 19 Country: United States Co-morbidity: none, healthy infants Socio-demographics: not reported. Ethnicity: 16 Black, 4 White, 6 Hispanic, 1 American Indian, 3 not classified
Interventions	Total number of intervention groups: 1 Control group description: wrapped in a receiving blanket and placed on their side in their respective bassinets Total number of participants randomized to each of your groups: intervention = 15; control = 15 (total = 30) Specific intervention: skin-to-skin contact Intervention details: the bed was adjusted to 45 degrees. The infant wore only a diaper and was positioned on mother. Two blankets were placed over the infant with the infant's heel accessible. Mothers locked their fingers and placed their hands over the blanket while applying slight pressure on the infant's back. Mother and infant were left alone for 10 to 15 minutes prior to procedure Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **Grimacing (brow bulge, eye squeeze and nasolabial furrowing)** 3 minutes post heel lance Upper limit: 1 (present) Lower limit: 0 (not present) Higher score = higher pain Continuous **Duration of cry** 3 minutes post heel lance Continuous
Notes	Adverse reactions: data collection was terminated for 4 control infants during the second minute and for 7 control infants during the third minute due to excessive crying Funding sources: National Institute of Mental Health; Maternal and Child Health Bureau Key conclusions of the study authors: skin-to-skin contact is an effective, easily implemented and safe intervention against pain in newborns experiencing the heelstick procedure

	Miscellaneous comments from the study authors: some mothers in the intervention group "spoke gently to their infants or made clicking sounds."

Risk of bias

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori

Gray 2000 (*Continued*)

Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	Low risk	18 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Did not report whether they were blinded or not.
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Greenberg 2002

Methods	Study design: randomized between groups Total study duration: unknown
Participants	Total number: 84 Setting: moderate sized hospital, nursery Diagnostic criteria: birth weight between 2500 and 4000 grams; less than or equal to 72 hours old; no congenital abnormalities; full-term Age: Minimum: age of birth Maximum: 72 hours Mean: 18.99 hours SD: 5.59 hours Sex: Males: 38 Females: 46 Country: United States Co-morbidity: none Socio-demographics: mothers were excluded from the study if there was a history of sub-

	stance abuse, or chronic or infectious diseases during pregnancy Ethnicity: 73.8% white; 10.7% Hispanic; 4.8% African American; 10.7% other
Interventions	Total number of intervention groups: 3 (we will only focus on 1 for review, as the other intervention group included sucrose) Control group description: routine care; no intervention offered during heelstick Total number of participants randomized to each of your groups: 21 in each of the 4 groups (total of 84) Specific intervention: water pacifier Intervention details: a Mini-Mam 0 to 6-month orthodontic pacifier moistened with water was gently held in the infant's mouth (without initiating any other contact) during 2 minutes pre procedure and 3 minutes post procedure Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **Duration of cry** From the start of the procedure to 25 minutes post procedure Upper limit: 1500 seconds (and higher) Lower limit: 0 seconds Higher score = longer cry duration, more pain Continuous
Notes	Adverse reactions: none Funding sources: American Cancer Society; California Division Fellowship Key conclusions of the study authors: no significant differences between water pacifier and typical care group were found

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history

Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	High risk	0 - Most of the measures were not reliable or sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	High risk	11 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	Investigator collected all data

Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Herrington 2007

Methods	Study design: cross-over Total study duration: 4 months (October 2006 to January 2007)
Participants	Total number: 11 (21 enrolled initially; 1 withdrew, 2 were excluded, 6 discharged, 1 excluded due to video difficulty) Setting: NICU Diagnostic criteria: premature Age: Minimum: 27 weeks Maximum: 33 weeks Mean: 31 weeks SD: not reported Sex: Males: 6 Females: 5 Country: United States Ethnicity: 57.5% of the infants were African American; 40% Caucasian; 2.5% Hispanic origin
Interventions	Total number of intervention groups: 1 Control group description: no intervention, PI did not directly touch the infant. Both hands were loosely fisted and placed behind the infant in close approximation to the head and buttocks, but not in contact with the infant Total number of participants randomized to each of your groups: 11 Specific intervention: gentle human touch Intervention details: gentle positioning support for the infant using warm human hands to contain the infant's body in a flexed position. PI placed both hands in isolette, placing the right hand behind the infant's head and shoulders in a cupped fashion. PI paused for 10 seconds to release tension. The position was held for the entire duration of the heelstick procedure and 2 minutes post heelstick procedure Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **Cry duration** Heelstick and 2 minutes post draw recovery Upper limit: (variable) Lower limit: 0 seconds Higher score = more pain Continuous

Notes	Adverse reactions: none Funding sources: none reported Key conclusions of the study authors: gentle human touch does not reduce cry duration in moderately premature infants receiving the heelstick procedure Miscellaneous comments by review authors: study was double-blind; participant information is not known

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid

Herrington 2007 (*Continued*)

Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	20 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Hill 2005

Methods	Study design: cross-over Total study duration: unknown
Participants	Total number: 12 Setting: NICU Diagnostic criteria: preterm Age: Minimum: 25 weeks Maximum: 34 weeks Mean: 30.9 weeks SD: 2.5 weeks

	Sex: Males: 6 Females: 6 Country: USA
Interventions	Total number of intervention groups: 1 Control group description: 1 Total number of participants randomized to each of your groups: 12 Specific intervention: facilitated tucking Intervention details: nurse provided care to the infant while the physical therapist supported the infant in a midline, tucked position Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **PIPP** Immediate, during procedure Upper limit: 21 Lower limit: 0 Continuous
Notes	Adverse reactions: none Funding sources: none stated. Key conclusions of the study authors: by incorporating facilitated tucking into routine care events, the stress levels of infants born preterm may be reduced

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history

Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	0 - Power calculations were not conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	21 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified

Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Hillgrove Stuart 2008

Methods	Study design: randomized between groups Total study duration: 11 months
Participants	Total number: 99 Setting: pediatrician's clinic Diagnostic criteria:health infants Age: Minimum: 12 weeks Maximum: 20 weeks Mean: 14.98 weeks SD: 2.88 weeks Sex: Males: 58 Females: 41 Country: Canada Co-morbidity: Socio-demographics: predominantly high Ethnicity: predominantly white
Interventions	Total number of intervention groups: 2 Control group description: no intervention Total number of participants randomized to each of your groups: 34 to control, 33 to RA-led distraction, 32 to parent-led distraction Specific intervention: RA-led distraction Intervention details: research assistant held toy and distracted infant Integrity of intervention: good Specific intervention #2 : parent-led distraction Intervention details: parent distracted infant throughout procedure Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **MBPS** 15 seconds to 1 min after needle Upper limit: (variable): 10 Lower limit: 0 seconds Higher score = more pain Continuous
Notes	Adverse reactions: none Funding sources: not stated. Key conclusions of the study authors: toy distraction does not work

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori

Hillgrove Stuart 2008 (*Continued*)

Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used
Total	Low risk	20 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Im 2008

Methods	Study design: randomized between group Total study duration: unknown
Participants	Total number: 99 Setting: hospital Diagnostic criteria: healthy neonates Age: Minimum: 38.8 weeks Maximum: 39.1 weeks Mean: approximately 273.4 days gestation SD: approximately 9.31 days gestation Sex: Males: 47 Females: 52 Country: South Korea Socio-demographics:unknown Ethnicity: Korean
Interventions	Total number of intervention groups: 2 Control group description: 1

	Total number of participants randomized to each of your groups: touch: 33; non-nutritive sucking: 33; control: 33 Specific intervention #1: Yakson touch Intervention details: nurse warmed hands to 34 degrees Celsius and approached neonate from behind. Nurse placed left hand under neonate's back. Nurse placed right hand on neonate's abdomen and caressed clockwise (approximately 4 cm in diameter) every 5 seconds Integrity of intervention: good Specific intervention #2: non-nutritive sucking Intervention details: neonates were given a small, short and hollow nipple packed with sterile gauze while gentle pressure was applied in the mouth Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIPS** Taken at one minute post heelstick Upper limit: 7 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: none reported Key conclusions of study authors: touch and non-nutritive sucking were found not to reduce pain in healthy neonates measured 1 minute after heelstick, as compared to the control group

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	High risk	0 - Insufficient information is reported to allow adequate comparisons to be made

Im 2008 (*Continued*)

Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	High risk	13 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified

Im 2008 (*Continued*)

Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Ipp 2004

Methods	Study design: randomized between groups Total study duration: unknown
Participants	Total number: 106 Setting: community pediatrics office Diagnostic criteria: healthy infants Age: Minimum: 2 months Maximum: 6 months Mean: 123 days SD: 51 days Sex: Males: 47 Females: 59 Country: Canada Co-morbidity: 60% of the males were circumcised.
Interventions	Total number of intervention groups: 1 (held by parent) Control group description: 1 (positioned supine) Total number of participants randomized to each of your groups: 50 in the control group, 56 in the intervention group Specific intervention: holding Intervention details: mothers were instructed to stand and hold the baby in a comfortable way (no specific instructions given) Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time when measured: **NFCS** (brow bulge, nasolabial furrowing, and eyes squeezed shut) Time: not mentioned Upper limit: 300% Lower limit: 0% High score = more pain Continuous **Crying** 180 seconds Upper limit: 180 seconds Lower limit: 0 More crying = more pain Continuous **VAS**

	15 seconds post needle Upper limit: 100 mm (maximum pain) Lower limit: 0 mm (no pain) High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: none reported. Key conclusions of the study authors: holding is not better than the supine positioning in reducing the post needle pain in healthy infants aged 2 to 6 months (based on crying and NFCS and VAS scores) Miscellaneous comments by review authors: it was not mentioned how long NFCS was coded for, though it was stated that the maximum was 180 seconds)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators

Ipp 2004 (*Continued*)

Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	19 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Jain 2006

Methods	Study design: randomized cross-over Total study duration: unknown
Participants	Total number: 23 Setting: hospital NICU Diagnostic criteria: premies who were not currently ventilated and not on analgesic for the last 48 hours, and did not undergo surgery in the last 2 weeks Age: Minimum: 28 weeks Maximum: 35 weeks Mean: 31.1 weeks SD: 1.9 weeks Sex: no information stated Country: Canada (Alberta)
Interventions	Total number of intervention groups: 1 Control group description: 1 Total number of participants randomized to each of your groups: 23 Specific intervention: massage Intervention details: the infant was wrapped and bundled for 5 minutes prior to heel lance. The heel (from toes to mid thigh) was massaged with gentle pressure using fingers and thumbs. The heel was then covered with a warm cloth for 2 to 3 minutes Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIPS** 5 minutes post heel lance High score = more pain Continuous
Notes	Adverse reactions: data for 3 infants was not included (they were transferred to another hospital) Funding sources: no mentioned. Key conclusions of the study authors: gentle massage prior to heel lance is helpful in reducing pain during heel lance in premature infants

Risk of bias

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed

Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis

Jain 2006 (*Continued*)

Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	19 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Johnston 1997

Methods	Study design: between groups Total study duration: unknown
Participants	Total number: 90 randomized, in the end there were 85 participants Setting: NICU Diagnostic criteria: premature sample Age: Minimum: 25 4/7 weeks Maximum: 34 2/7 weeks Mean: 31 weeks SD: not reported Sex: Males: 42 Females: 43 Country: Canada Co-morbidity: no congenital abnormalities; Apgar scores greater than 3 at 5 minutes Socio-demographics: not reported Ethnicity: not reported
Interventions	Total number of intervention groups: 1 Control group description: baby was handled and repositioned to the side, and 0.05 ml of water was placed on the tongue's surface just prior to starting the heel lance procedure Total number of participants randomized to each group: 24 in rocking condition 20 in control condition Specific intervention: rocking Intervention details: the infant was repositioned to be on the side or in the supine position and was swaddled in a blanket. An air mattress was placed beneath them, that pumped air into the mattress at a rate of 12 cycles per minute. The swaddled baby was left on the oscillating mattress for 15 minutes prior to the heelstick. The infant also received 0.05 ml

	of water Integrity of intervention: good
Outcomes	Pain outcomes, time points when measured, and scale limits: **NFCS** Baseline, 30 seconds post needle, 60 seconds post needle, 90 seconds post needle Upper limit: 1 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: FRSQ, NIH Key conclusions: no difference between rocking and control. Sucrose is good for pain, though rocking is not

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants

Johnston 1997 (*Continued*)

Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	16 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Johnston 2003

Methods	Study design: cross-over design Total study duration: 14 months
Participants	Total number: 74 Setting: NICU Diagnostic criteria: premature infants Age Minimum: 32 0/7 weeks Maximum: 36 6/7 weeks Mean: 33.7 weeks SD: 1.1 weeks Sex Males: 41 Females: 33 Country: Canada
Interventions	Total number of intervention groups: 1 Control group description: the infant was placed in isolette in prone position and swaddled for 30 minutes pre heelstick Total number of participants randomized to each of your groups: 74 Specific interventions: kangaroo care Intervention details: the diaper clad infant was held upright at a 60 degree angle between the mother's breasts, skin-to-skin, with a blanket over them for 30 minutes pre heelstick. The mother's hands were clasped behind the neonate's back. The mother was instructed not to vocalize or touch the neonate's head with her face during filming Integrity of intervention: good
Outcomes	Pain outcomes, time points when measured, and scale limits: **PIPP** 30, 60, 90, 120 seconds post heel lancing Upper limit: 21 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: Canadian Institute of Health Research Key conclusions of study authors: kangaroo care works for preterm infants (lower PIPP scores at 30, 60, 90 seconds), but no difference at 120 seconds Miscellaneous comments from the study authors: nurses reported it was more difficult to obtain blood for kangaroo care infants due to neonate's positioning, although the effect would have been opposite otherwise

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**

Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	Low risk	1 - Most of the measures are valid
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial

Johnston 2003 (*Continued*)

Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	19 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	Low risk	NFCS coders were blinded
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Johnston 2007

Methods	Study design: cross-over Total study duration:
Participants	Total number: 39 Setting: NICU Diagnostic criteria: preterm infants Age: Minimum: 32 weeks Maximum: 34 5/7 weeks Mean: 33 weeks Std. 0.89 weeks Sex: Males: 57% males Females:43% females Country: Canada
Interventions	Total number of intervention groups: 1 Control group description: no intervention Total number of participants randomized to each of your groups: 20 Specific intervention: mother's voice played during heel lance Intervention details: recording of mothers voice either singing, talking, or reading a nursery rhyme was played 1 minute before the procedure began and until heart rate and O2 sat returned to baseline

	Integrity of intervention: good
Outcomes	Pain outcomes, time points when measured, and scale limits: 1) **PIPP** 30 seconds Upper limit: 21 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: Canadian Institute of Health Research Key conclusions of the authors: playing maternal voice pre and during heel lance has no significant impact on premie pain Miscellaneous comments by review authors: no means, only F provided for PIPP Percentages for NFCS

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants

Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	High risk	0 - There is inadequate reported of the summary statistics, i.e. the means, standard deviations were not reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	Low risk	NFCS coders were blind to the purpose of the study
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Kashaninia 2008

Methods	Study design: between groups Total study duration: 2 months
Participants	Total number: 100 Setting: hospital Diagnostic criteria: normative sample Age: Minimum: unknown Maximum: unknown Mean: 39.24 weeks SD: 1.44 weeks Sex: Males: 22 Females: 28 Country: Iran Co-morbidity: none
Interventions	Total number of intervention groups: 1 Control group description: infants were brought to a quiet room, were repositioned and left for 10 minutes Total number of participants randomized to each of your groups: 50 randomized to each group Specific intervention: kangaroo care Intervention details: diapered infant was placed so that there was skin-to-skin contact through the open gown. Chair was at a 45 degree angle. They received blankets that were placed over the infant's back. Mother locked fingers and placed them under infant's back, and applied slight pressure. Mother was instructed not to rub, speak or jiggle the infant. The infant was left to settle for 10 minutes pre needle. Kangaroo care continued during the needle Integrity of intervention: good
Outcomes	Pain outcomes, time points when measured, and scale limits: **NIPS** Measured for 1 minute immediately after the needle Upper limit: 7 Lower limit: 0 High score = more pain Categorical scale: 0 to 2: no pain to mild pain 3 to 4: moderate pain > 4: severe pain
Notes	Adverse reactions: none Funding sources: none mentioned Key conclusions of study authors: kangaroo care works for typical newborn neonates with needle for vitamin K
Risk of bias	

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori

Kashaninia 2008 (*Continued*)

Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	16 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Kostandy 2008

Methods	Study design: randomized cross-over Total study duration: unknown
Participants	Total number: 10 Setting: level II NICU Diagnostic criteria: premature Age: Minimum: 2 days Maximum: 9 days Mean: 31.4 SD: unknown Sex: Males: 50% Females: 50% Country: USA Co-morbidity: none Socio-demographics: 70% of the mothers were married; 30% were employed full-time,

	30% were employed part-time, 20% were students, 20% were homemakers; 10% had completed graduate school, 40% had completed college, 20% had some college and 30% had completed high school Ethnicity: 90% white
Interventions	Total intervention groups: 2 Control group description: no active control Total number of participants randomized to each of your groups: 10 Specific intervention #1: kangaroo care Intervention details: skin-to-skin, chest-to-chest, upright placement of the infant wearing only a diaper between maternal breasts, with receiving blanket over the infant's back Integrity of intervention: not needed Specific intervention #2:incubator Intervention details: infant was nested between rolled blankets on a mattress of 30 to 40 degree incline Integrity of intervention: not needed
Outcomes	Pain outcomes, scale limits, and time points when measured: **Crying time** Summing the total number of seconds of inaudible crying (Anderson Behavioral State Scoring System) and the total number of seconds of audible crying
Notes	Adverse reactions: none Funding sources: 26 subjects were needed to detect moderate difference in crying time, however, funding only permitted 10 subjects (pilot study), National Institutes of Health, Association of Women's Health, Obstetrics, and Neonatal Nurses Grant Key conclusions of study authors: kangaroo care diminishes crying time

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant

Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	High risk	0 - Either equivalence of groups is not reported or there is evidence of non-equivalence
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness

Kostandy 2008 (*Continued*)

Total	Low risk	16 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Kozub 2001

Methods	Study design: randomized cross-over Total study duration: winter of 2000
Participants	Total number: 20 Setting: pulmonary clinic (hospital) Diagnostic criteria: premature born infants (but older at time of the study), receiving Synagis injections to prevent pulmonary infections Age: Minimum: 1 month Maximum: 15 months Mean: 6.58 months SD: 4.1 months Sex: Males: 12 Females: 8 Country: USA
Interventions	Total intervention groups: 1 Control group description: 1 (mimic intervention) Total number of participants randomized to each of your groups: 20 Specific intervention: therapeutic touch Intervention details: Krieger method: centering, assessment, unruffing and direction/modulation of energy Centering: focus attention Assessment: detection of irregular energy areas (chakras) Unruffing: smoothing at energy field Direction/modulation of energy: energy transfer (change flow) and using visualization and intentionality to redistribute energy flow Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **PIPP** 30 seconds post injection Upper limit: 21 Lower limit: 0 High score = more distress

	Continuous	
Notes	Adverse reactions: none Funding sources: none reported Key conclusions of study authors: therapeutic touch does not reduce immediate (30 seconds) pain during injections in infants aged 1 to 15 months	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	Low risk	2 - Documented evidence that the CONSORT guidelines have been followed
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study

Kozub 2001 (*Continued*)

Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	22 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Liaw 2010

Methods	Study design: between groups Total study duration: 9 months
Participants	Total number: 104 Setting: NICU in Taiwan Diagnostic criteria: healthy, preterm infants, between 3 and 28 days post birth Age: 27 to 37 weeks Minimum: 3 days Maximum: 19 days Mean: 6.48 days SD: 3.11 days

	Sex: Males: 54 Females: 50 Country: Taiwan
Interventions	Total intervention groups: 1 Control group description: routine comfort (gentle touching) without non-nutritive sucking Total number of participants randomized to each of your groups: 52 Specific intervention: non-nutritive sucking Intervention details: pacifier was given to infant to suck before touching the foot to initiate the heelstick Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **PIPP** Every minute for before (for 3 minutes), during, and after (for 10 minutes) the procedure Upper limit: 21 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none reported Funding sources: none reported Key conclusions of study authors: non-nutritive sucking reduces pain, especially mild to moderate pain and behavioral responses

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history

Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	22 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)

Blinding	High risk	None specified
Selective outcome reporting	Low risk	None specified
Other potential sources of bias	Low risk	None specified

Liu 2010

Methods	Study design: randomized control trial, between groups (parallel-group controlled trial) Total study duration: 5 months (June to October 2006)
Participants	Total number: 105 Setting: newborn special care nursery and in a newborn baby room of a teaching hospital in northern Taiwan Diagnostic criteria: neonates were greater than or equal to 32 weeks gestational age, medically stable, scheduled to undergo a newborn screening procedure within 24 hours to 7 days, Apgar scores greater than 7 at 1 and 5 minutes, with no crying in the 5 minutes prevenipuncture Age: Minimum: unknown Maximum: unknown Mean: 3.03 days SD: 1.06 days Sex: Males: 48 Females: 57 Country: Taiwan
Interventions	Total intervention groups: 1 Control group description: no intervention provided prior to venipuncture. Infants were taken to a quiet, individual room and placed on a heated radial warmer, naked except for their diaper. Infants were observed for 2 minutes to collect baseline data Total number of participants randomized to each of your groups: 35 in each group (70 in total) Specific intervention: non-nutritive sucking Intervention details: each neonate was assisted to suck on a pacifier for 2 minutes presterilization. The pacifier was continuously provided until 2 minutes post needle Integrity of intervention: good
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIPS** Coded at 1-minute intervals for 2 minutes in preparation, venipuncture, and recovery phases Upper limit: 7 Lower limit: 0 High score = more pain Continuous

Notes	Adverse reactions: none reported Funding sources: no financial sources that might pose a conflict of interest declared Key conclusions of study authors: non-nutritive sucking can effectively decrease the level of pain

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	High risk	0 - There is no clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	High risk	0 - Either equivalence of groups is not reported or there is evidence of non-equivalence
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	Low risk	1 - An adequate method is reported that removes the potential biases of investigators
Measurement bias	Low risk	1 - A convincing effort to reduce bias in outcomes measurement is reported
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study

Liu 2010 (*Continued*)

Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	18 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	Low risk	Nurses who coded NIPS were blind to group
Selective outcome reporting	Low risk	None specified
Other potential sources of bias	Low risk	None specified

Morrow 2010

Methods	Study design: between groups Total study duration: 3 months (February to April 2008)
Participants	Total number: 42 Setting: tertiary hospital serving a major metropolitan area Diagnostic criteria: infants scoring in the high-intermediate or high-risk zone by TcB (Transcutaneous Bilirubin) measurement receiving a total serum bilirubin (TSB) evaluation; full-term neonates Age: greater than or equal to 37 weeks gestation Minimum: not reported Maximum: not reported Mean: 39.13 weeks gestation

	SD: 1.84 days Sex: Males: 18 Females: 21 Country: United States
Interventions	Total intervention groups: 1 Control group description: infants were placed in a standard position (supine while lying in a crib) during a sample collection. The crib was elevated to a 30 degree angle and the leg was elevated during the heel lance Total number of participants randomized to each of your groups: 42 Specific intervention: swaddling and holding upright Intervention details: infants were swaddled and held upright at a 90 degree angle, with one leg exposed Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIPS** Coded 1 minute pre heel lance (immediately before) and 1 minute post heel lance (immediately after) Upper limit: 7 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none reported Funding sources: no funding sources. Key conclusions of study authors: swaddling infants while holding them in an upright position was superior for pain relief during heel lance procedures when compared with a standard position technique

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	High risk	0 - There is no clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias

Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	1 - A convincing method of randomization is reported but this did not involve an independent person
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness

Morrow 2010 (*Continued*)

Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	The nurses who performed the heel lance also rated NIPS scores
Selective outcome reporting	Low risk	None specified
Other potential sources of bias	Low risk	None specified

Sizun 2002

Methods	Study design: randomized cross-over Total study duration: unknown
Participants	Total number: 19 Setting: hospital NICU Diagnostic criteria: preterm neonates, younger than 32 weeks Age: Minimum: unknown Maximum: unknown Mean: 29 weeks SD: 1.8 weeks Sex: Males: 10 Females: 19 Country: France
Interventions	Total number of intervention groups: 1 Control group description: no protection from light, supine, no supportive bedding, no individualized attention Total number of participants randomized to each of your groups: 19 Specific intervention #1: developmental care Intervention details: decreased light, decreased noise, lateral posture, feet contacting supportive bedding, opportunities for grasping, offered motor support by hard swaddling, grasping, support for state transition Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **PIPP** Immediate Upper limit: 21 Lower limit: 0 Higher score = more pain Continuous

Notes	Adverse reactions: none Funding sources: PHRC 1998 grant Key conclusions of study authors: developmental care works on reactivity	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	High risk	0 - There is no clear evidence that the inclusion/exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	High risk	0 - Randomization is mentioned but there is not an adequate description of the methods used
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study

Sizun 2002 (*Continued*)

Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	14 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Stevens 1999

Methods	Study design: randomized cross-over Total study duration: 15 months
Participants	Total number: 122 Setting: NICU Diagnostic criteria: preterm infants Age: Minimum: 27 weeks Maximum: 31 weeks Mean: 28.6 weeks SD: 0.15 weeks

	Sex: Males: 70 males Females: 52 females Country: Canada Co-morbidity: none Socio-demographics: not reported Ethnicity: 78 white, 12 black, 17 Hispanic, 15 Asian
Interventions	Total number of intervention groups: 2 Control group description: in a Snuggle-Up device for 30 minutes prior to heel lance procedure. Infant was side-lying or in the supine position in the Snuggle-Up Total number of participants randomized to each of your groups: 122 Specific intervention #1: prone positioning Intervention details: infants were positioned prone in Snuggle-Up, with knees flexed, arms in, arms close to midline and left foot free Integrity of intervention: good Specific Intervention #2: pacifier with water Intervention details: infants were given a pacifier dipped in water and positioned in Snuggle-Up for 7 minutes pre lance Integrity of intervention: good
Outcomes	Pain outcomes, time points when measured, and scale limits: **PIPP** Immediately (stick 15 seconds) 5 minutes after Upper limit: 21 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: NIH, NINR, NIHPCRC Key conclusions of study authors: pacifier with water was the most effective method in reducing pain

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT

Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	Low risk	2 - A convincing method for generating a random allocation sequence is reported that used an independent person not involved in enrolment or allocation of participants
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis

Stevens 1999 (*Continued*)

Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	High risk	There were no other potential sources of bias reported

Ward-Larson 2004

Methods	Study design: cross-over Total study duration: not noted
Participants	Total number: 40 Setting: NICU Diagnostic criteria: very preterm infants Age: Minimum: 23 weeks Maximum: 32 weeks Mean: 27.313 weeks SD: 2.430 weeks Sex: Males: 22 Females: 18 Country: USA Co-morbidity: none Socio-demographics: none listed Ethnicity: 23 African American, 16 Caucasian, 1 Hispanic
Interventions	Total number of intervention groups: 1 Control group description: infants were put in a Snuggle-Up (provides containment). No hands were on the infants Total number of participants randomized to each of your groups: 40 Specific intervention: facilitated tucking Intervention details: infants were put in a Snuggle-Up and turned to their sides. Their backs were curled gently, legs flexed at a 90 degree angle and their shoulders were brought to midline. They remained like this for suctioning and for 30 seconds after. There was a 2 to 4-hour washout period between procedures Integrity of intervention: good

Outcomes	Pain outcomes, time points when measured, and scale limits: **PIPP** Measured for 30 seconds after suctioning Upper limit: 21 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none Funding sources: not mentioned Key conclusions of study authors: facilitated tucking works

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	High risk	0 - Randomization is mentioned but there is not an adequate description of the methods used
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist

Ward-Larson 2004 (*Continued*)

Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	Low risk	1 - The analysis included an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	Low risk	Investigator did the coding

Whipple 2004

Methods	Study design: between groups Total study duration: unknown
Participants	Total number: 60 Setting: NICU-SCN Level III of Hospital in Atlanta, Georgia Diagnostic criteria: low birth weight (< 2500 g), hospitalized in NICU, preterm Age:

	Minimum: 24.43 weeks Maximum: 35.43 weeks Mean: 31.17 weeks SD: not reported for whole sample Sex: Males: 30 Females: 30 Country: USA
Interventions	Total number of intervention groups: 2 Control group description: no contact, infants were not provided with a pacifier or music listening opportunities at any point during heelstick procedure. Standard care and pain management procedures (such as swaddling, cuddling, and sucrose) were not limited to infants of any of the groups Total number of participants randomized to each of your groups: 20 in each group (total of 60) Specific intervention #1: Sondrez PAL system-music contingent on sucking Intervention details: intervention began 3 minutes prior to heelstick and continued throughout the procedure until approximately 3 minutes after blood collection. PAL played a lullaby contingent on the sucking of the infant Integrity of Intervention: satisfactory Specific intervention #2: pacifier provided, no music Intervention details: same procedures as the intervention #1, except that the infants did not receive music reinforcement for sucking Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **Behavioral state** (scored based on the "Assessment of Premature Infant's Behavior") 1 second intervals and 1 minute pre needle, last minute pre needle, 1 minute during needle, 1 minute post needle, last minute post needle Upper limit: 254 Lower limit: 15 High score = high pain Continuous **Stress level** 1 second intervals and 1 minute pre-needle, last minute pre-needle, 1 minute during needle, 1 minute post needle, last minute post needle Upper limit: 837 Lower limit: 54 High score = high pain Continuous
Notes	Adverse reactions: none Funding sources: none mentioned Key conclusions of study authors: music reinforced NNS, effectively lowers behavioral states and stress levels in low birthweight, preterm infants undergoing heelstick
Risk of bias	

Whipple 2004 (*Continued*)

Bias	Authors' judgement	Support for judgement
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/ exclusion criteria were met
Attrition	High risk	0 - There is no documented evidence or insufficient evidence reported of how attrition was dealt with
Rates of attrition	High risk	0 - There is insufficient evidence that differential rates of attrition have not resulted in significant bias
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data
Randomization	High risk	0 - Randomization is mentioned but there is not an adequate description of the methods used
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	1 - Most of the outcome measures are justified
Validity of outcomes for context	High risk	0 - Most or all of the measures are not valid given the content of the particular study
Reliability and sensitivity	Low risk	1 - Most of the measures were reliable and sensitive to change
Power calculation	High risk	0 - Power calculations were not conducted a priori

Whipple 2004 (*Continued*)

Sufficient sample	High risk	0 - A sufficient sample size based on the power calculation was not obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	High risk	0 - A poor control group has been used that merely control for the duration of time
Total	High risk	8 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	High risk	There was no selective outcome reporting
Other potential sources of bias	Low risk	Were randomized prior to setting the date for the heelstick. Therefore, it was known what condition they were in before the heel procedure was completed

Yilmaz 2010

Methods	Study design: between groups, RCT prospective Total study duration: 12 months
Participants	Total number: 120 Setting: hospital Diagnostic criteria: low birth weight (< 2500 g), Apgar score greater than or equal to 8, fed at least 30 minutes before procedure, no signs of disease or congenital anomalies, C-Section delivery, blood sampling at the first attempt, mothers holding baby in their arms during the procedure Age: 37 to 42 weeks (gestational age) Minimum: 37 weeks GA Maximum: 42 weeks GA Mean: 3.33 days SD: 1.24 Sex: Males: 61 Females: 59 (49.2%)

	Country: Turkey
Interventions	Total number of intervention groups: 1 Control group description: baby in mother's lap, no interventions Total number of participants randomized to each of your groups: 30 Specific intervention #1: pacifier Intervention details: babies were given a pacifier Integrity of intervention: satisfactory
Outcomes	Pain outcomes, scale limits, and time points when measured: **NIPS** Coded for 2 minutes before, 5 minutes during, and 3 minutes after injections Upper limit: 7 Lower limit: 0 High score = more pain Continuous
Notes	Adverse reactions: none mentioned Funding sources: none mentioned Key conclusions of study authors: pacifying shortened the length of crying time and the rate of behavioral responses to pain compared to the control group

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Sample criteria	Low risk	1 - The inclusion/exclusion criteria are clearly specified for the sample and there is evidence of adherence to the criteria
Evidence that the criteria has been met	Low risk	1 - There is clear evidence that the inclusion/exclusion criteria were met
Attrition	Low risk	1 - A reasonable account of how attrition was dealt with is given, but no specific reference to CONSORT
Rates of attrition	Low risk	1 - There is evidence that any differential rates of attrition were not statistically significant
Sample characteristics	Low risk	1 - There is a good description of the sample in the trial detailing areas such as demographic details treatment history
Group equivalence	Low risk	1 - There is evidence that the groups are broadly equivalent shown by testing or examination of reported data

Randomization	High risk	0 - Randomization is mentioned but there is not an adequate description of the methods used
Allocation bias	High risk	0 - There is not an adequate description of attempts to deal with potential allocation bias
Measurement bias	High risk	0 - Efforts to reduce measurement bias are not reported or are insufficient, e.g. outcomes collected by therapist
Justification of outcomes	Low risk	2 - All of the outcomes measures are justified
Validity of outcomes for context	Low risk	2 - All of the outcome measures are valid given the context of the study
Reliability and sensitivity	Low risk	2 - All the outcome measures chosen were shown to be reliable and sensitive to change
Power calculation	Low risk	1 - Power calculations were conducted a priori
Sufficient sample	Low risk	1 - A sufficient sample size based on the power calculation was obtained
Planned analysis	Low risk	1 - The data analysis was adequately planned to assess the hypothesis and aims of the trial
Statistics reporting	Low risk	1 - There is adequate reported of the summary statistics, i.e. the means, standard deviations were reported
Intention-to-treat	High risk	0 - The analysis did not include an intention-to-treat analysis
Adequate control group	Low risk	1 - An active alternative treatment group has been used that is well matched in terms of structural features of the treatment and its meaningfulness
Total	Low risk	17 - Total score on quality rating (Yates, Morley, Eccleston, Williams, 2005)
Blinding	High risk	None specified
Selective outcome reporting	Low risk	None specified

Other potential sources of bias	Low risk	None specified

CONSORT: Consolidated Standards of Reporting Trials
DAN: Douleur Aiguë Nouveau-né
EBI: Environmental and Behavioral Interventions
EDIN: Échelle Douleur Inconfort Nouveau-Né (neonatal pain and discomfort scale)
EMLA: Eutectic Mixture of Local Anesthetics
ER: emergency room
FLACC: Face Legs Arms Cry Consolability Scale
FRSQ: Fonds de recherche en santé du Québec
GA: gestational age
MAISD: Measure of Adult and Infant Soothing and Distress
MBPS: Modified Behavioral Pain Scale
NFCS: Neonatal Facial Coding System
NICU: neonatal intensive care unit
NIDCAP: Newborn Individualized Developmental Care and Assessment Program
NIH: National Institute of Health
NIH PCRC: National Institute of Health Pediatric Clinical Research Center
NINR: National Institute of Nursing Research
NIPS: Neonatal Infant Pain Scale
NNS: Non-nutritive sucking
PHRC: *French* Government *Funding* Program for Clinical Research
PI: Principal Investigator
PIPP: Premature Infant Pain Profile
RA: research assistant
RCT: randomized controlled trial
s: seconds
SES: socioeconomic status
SESEP: Société d'Études et de Soins pour les Enfants Paralysés et Malformés
TcB: Transcutaneous Bilirubin
VAS: visual analog scale

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

Study	Reason for exclusion
Abedin 2008	Wrong age or age could not be separated
Aguirre 2008	Control group is an active control group Included in 'Summary of treatment effect' section
Al-Bekaa 2003	Wrong age or age could not be separated
Aslanabadi 2008	Wrong age or age could not be separated

(*Continued*)

Axelin 2010	Inappropriate intervention
Bellieni 2003	Inappropriate intervention
Bellieni 2007	Inappropriate intervention
Boots 2010	Inappropriate intervention
Bueno 2010	Control group was an active control Included in 'Summary of treatment effect' section
Campbell-Yeo 2009	Randomized controlled trial in process
Campos 1989	Control group was an active control Included in 'Summary of treatment effect' section
Carlson 2000	Wrong age or age could not be separated
Cignacco 2008	Control group was an active control Included in 'Summary of treatment effect' section
Cologna 1999	Not pain or pain is not measured immediately
Cong 2009	No behavioral pain outcome reported Included in 'Summary of treatment effect' section
D'Agostino 2008	Wrong age or age could not be separated
Dahlquist 2002	Wrong age or age could not be separated
Diego 2009	Control group is an active control group Included in 'Summary of treatment effect' section
Dilen 2010	Inappropriate intervention
Drago 2009	Wrong age or age could not be separated
Duncan 2004	Not pain r pain is not measured immediately.
El-Naggar 2010	Inappropriate intervention
Favara-Scacco 2001	Wrong age or age could not be separated
Felt 2000	Not a randomized controlled trial.
Gedaly-Duff 1992	Wrong age or age could not be separated

(*Continued*)

Gold 2006	Wrong age or age could not be separated
Gonzalez 1989	Wrong age or age could not be separated
Gonzalez 1993	Wrong age or age could not be separated
Goubet 2003	Control group was an active control Included in 'Summary of treatment effect' section
Goubet 2007	Control group was an active control Included in 'Summary of treatment effect' section
Greenberg 1997	Student work later published and included in the review
Grunau 2004	No control group Included in 'Summary of treatment effect' section
Hanson 2010	Inappropriate intervention
Harrison 2000	This study did not use a painful stimulus
Hatem 2006	Wrong age or age could not be separated
He 2010	Inappropriate intervention
Heden 2009	Wrong age or age could not be separated
Hoffman 2006	Wrong age or age could not be separated
Holsti 2005	Inappropriate intervention
Hsu 1995	Not pain or pain not measured immediately
Huang 1999	Not pain or pain not measured immediately
Huang 2004	Control group is an active control group Included in 'Summary of treatment effect' section
Ida 2008	Wrong age or age could not be separated
Ipp 2007	Inappropriate intervention
Ipp 2009	No control group Included in 'Summary of treatment effect' section
Jackson 2008	Wrong age or age could not be separated

(*Continued*)

Jo 2007	Wrong age or age could not be separated
Johnston 2007b	Not pain or pain not measured immediately
Johnston 2008a	Control group is an active control group Included in 'Summary of treatment effect' section
Johnston 2009	Control group is an active control group Included in 'Summary of treatment effect' section
Kivijärvi 2008	Wrong age or age could not be separated
Koivusalo 2009	Wrong age or age could not be separated
Leclair 2007	Not pain or pain not measured immediately
Li 2007	Wrong age or age could not be separated
Ludington-Hoe 2005	Control group was an active control Included in 'Summary of treatment effect' section
MacLaren 2005	Wrong age or age could not be separated
Manne 1990	Wrong age or age could not be separated
Marchisotti 2007	Wrong age or age could not be separated
Marec-Berard 2009	Wrong age or age could not be separated
Marin Gabriel 2010	Not pain or pain is not measured immediately
McCarthy 2010	Wrong age or age could not be separated
Michel 2008	Wrong age or age could not be separated
Morelius 2009	Not a randomized controlled trial Included in 'Summary of treatment effect' section
Mucignat 2004	Inappropriate intervention
Okan 2010	No means and standard deviations reported Included in 'Summary of treatment effect' section
Ozdogan 2010	Inappropriate intervention
Park 2006	Could not contact author

(*Continued*)

Phipps 2005	Wrong age or age could not be separated
Rattaz 2005	No control group Included in 'Summary of treatment effect' section
Reichel 2007	Wrong age or age could not be separated
Salmon 2006	Wrong age or age could not be separated
Shapiro 2007	Wrong age or age could not be separated
Slater 2010	Inappropriate intervention
Slifer 2009	Wrong age or age could not be separated
Sparks 2007	Wrong age or age could not be separated
Sundararajan 2007	Wrong age or age could not be separated
Tanabe 2002	Wrong age or age could not be separated
Vignochi 2010	Not pain or pain not measured immediately
Vivancos 2010	Not a randomized controlled trial Included in 'Summary of treatment effect' section
Weissman 2009	Not a randomized controlled trial Included in 'Summary of treatment effect' section
Wisdorf-Houtkooper 1997	Inappropriate intervention
Zeltzer 1991	Wrong age or age could not be separated

DATA AND ANALYSES

Comparison 1. Preterm X kangaroo care X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	5	236	Std. Mean Difference (Random, 95% CI)	-1.12 [-2.04, -0.21]

Comparison 2. Preterm X kangaroo care X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	5	222	Std. Mean Difference (Random, 95% CI)	-0.77 [-1.50, -0.03]

Comparison 3. Neonate X kangaroo care X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	2	420	Std. Mean Difference (Random, 95% CI)	-0.89 [-2.89, 1.10]

Comparison 4. Neonate X kangaroo care X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	2	343	Std. Mean Difference (Random, 95% CI)	-0.66 [-1.73, 0.42]

Comparison 5. Preterm X swaddling/tucking-related X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	6	261	Std. Mean Difference (Random, 95% CI)	-0.97 [-1.63, -0.31]

Comparison 6. Preterm X swaddling/tucking-related X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	3	65	Std. Mean Difference (Random, 95% CI)	-0.75 [-1.14, -0.36]

Comparison 7. Neonate X swaddling/tucking X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress	1	42	Std. Mean Difference (Random, 95% CI)	-1.26 [-1.92, -0.60]

Comparison 8. Preterm X non-nutritive sucking-related X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	5	305	Std. Mean Difference (Random, 95% CI)	-0.42 [-0.68, -0.15]

Comparison 9. Preterm X non-nutritive sucking-related X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	4	226	Std. Mean Difference (Random, 95% CI)	-0.38 [-0.59, -0.17]

Comparison 10. Neonate X non-nutritive sucking-related X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	4	220	Std. Mean Difference (Random, 95% CI)	-1.45 [-2.34, -0.57]

Comparison 11. Neonate X non-nutritive sucking-related X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	7	325	Std. Mean Difference (Random, 95% CI)	-0.90 [-1.54, -0.25]

Comparison 12. Older infants X non-nutritive sucking-related X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	41	Std. Mean Difference (Random, 95% CI)	-0.89 [-1.53, -0.25]

Comparison 13. Preterm X swallowing water X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	36	Std. Mean Difference (Random, 95% CI)	-0.24 [-0.71, 0.23]

Comparison 14. Preterm X swallowing water X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	36	Std. Mean Difference (Random, 95% CI)	-0.23 [-0.70, 0.24]

Comparison 15. Neonate X swallowing water X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	50	Std. Mean Difference (Random, 95% CI)	0.11 [-0.45, 0.66]

Comparison 16. Neonate X swallowing water X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	34	Std. Mean Difference (Random, 95% CI)	-0.53 [-1.21, 0.16]

Comparison 17. Older infants X swallowing water X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	30	Std. Mean Difference (Random, 95% CI)	0.0 [-0.72, 0.72]

Comparison 18. Neonate X rocking/holding X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	2	131	Std. Mean Difference (Random, 95% CI)	-0.33 [-1.05, 0.39]

Comparison 19. Neonate X rocking/holding X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	2	81	Std. Mean Difference (Random, 95% CI)	-0.75 [-1.20, -0.30]

Comparison 20. Older infants X rocking/holding X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	106	Std. Mean Difference (Random, 95% CI)	0.23 [-0.15, 0.62]

Comparison 21. Preterm X simulated rocking + water X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	44	Std. Mean Difference (Random, 95% CI)	0.0 [-0.59, 0.59]

Comparison 22. Preterm X touch or massage-related X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	2	34	Std. Mean Difference (Random, 95% CI)	-0.71 [-2.33, 0.90]

Comparison 23. Neonate X touch or massage-related X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	40	Std. Mean Difference (Random, 95% CI)	-0.3 [-0.92, 0.32]

Comparison 24. Neonate X touch or massage-related X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	66	Std. Mean Difference (Random, 95% CI)	-0.24 [-0.73, 0.24]

Comparison 25. Older infants X touch or massage-related X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	20	Std. Mean Difference (Random, 95% CI)	-0.21 [-0.84, 0.41]

Comparison 26. Preterm X environment modification X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	2	64	Std. Mean Difference (Random, 95% CI)	-6.44 [-17.13, 4.26]

Comparison 27. Preterm X environment modification X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	45	Std. Mean Difference (Random, 95% CI)	-4.01 [-5.26, -2.77]

Comparison 28. Older infants X toy distraction X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	3	259	Std. Mean Difference (Random, 95% CI)	-0.10 [-0.35, 0.14]

Comparison 29. Older infants X toy distraction X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	133	Std. Mean Difference (Random, 95% CI)	-0.08 [-0.50, 0.33]

Comparison 30. Older infants X video distraction X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	90	Std. Mean Difference (Random, 95% CI)	-0.70 [-1.13, -0.27]

Comparison 31. Older infants X video distraction X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	126	Std. Mean Difference (Random, 95% CI)	-0.84 [-1.20, -0.47]

Comparison 32. Older infant X structured parent involvement X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	3	209	Std. Mean Difference (Random, 95% CI)	-0.26 [-0.70, 0.17]

Comparison 33. Older infants X structured parent involvement X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	288	Std. Mean Difference (Random, 95% CI)	0.02 [-0.21, 0.25]

Comparison 34. Preterm X mother's voice X reactivity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress reactivity	1	19	Std. Mean Difference (Random, 95% CI)	-0.29 [-0.94, 0.35]

Comparison 35. Older infants X parent present X immediate regulation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Pain-related distress regulation	1	278	Std. Mean Difference (Random, 95% CI)	-0.00 [-0.24, 0.23]

Analysis 1.1. Comparison 1 Preterm X kangaroo care X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 1 Preterm X kangaroo care X reactivity

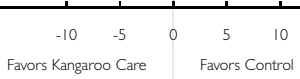
Outcome: 1 Pain-related distress reactivity

Study or subgroup	Kangaroo Care N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Ferber 2008	29	0	-10.362 (1.942)		4.6 %	-10.36 [-14.17, -6.56]
Kostandy 2008	5	0	-0.338 (0.65)		15.7 %	-0.34 [-1.61, 0.94]
Kostandy 2008	5	0	-0.151 (0.636)		15.9 %	-0.15 [-1.40, 1.10]
de Sousa 2008	31	33	-1.803 (0.297)		20.7 %	-1.80 [-2.39, -1.22]
Castral 2008	31	28	-0.613 (0.267)		21.1 %	-0.61 [-1.14, -0.09]
Johnston 2003	74	0	-0.31 (0.168)		22.0 %	-0.31 [-0.64, 0.02]
Total (95% CI)					**100.0 %**	**-1.12 [-2.04, -0.21]**

Heterogeneity: $Tau^2 = 0.96$; $Chi^2 = 45.02$, df = 5 (P<0.00001); $I^2 = 89\%$

Test for overall effect: Z = 2.41 (P = 0.016)

Test for subgroup differences: Not applicable



Analysis 2.1. Comparison 2 Preterm X kangaroo care X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 2 Preterm X kangaroo care X immediate regulation

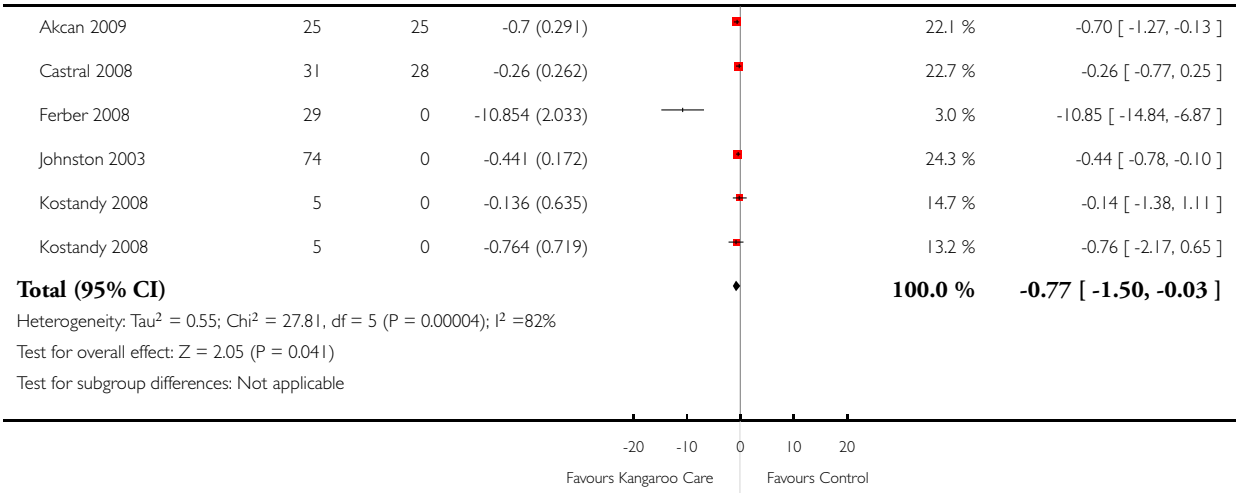
Outcome: 1 Pain-related distress regulation

Study or subgroup	Kangaroo Care N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Akcan 2009	25	25	-0.7 (0.291)		22.1 %	-0.70 [-1.27, -0.13]
Castral 2008	31	28	-0.26 (0.262)		22.7 %	-0.26 [-0.77, 0.25]
Ferber 2008	29	0	-10.854 (2.033)		3.0 %	-10.85 [-14.84, -6.87]
Johnston 2003	74	0	-0.441 (0.172)		24.3 %	-0.44 [-0.78, -0.10]
Kostandy 2008	5	0	-0.136 (0.635)		14.7 %	-0.14 [-1.38, 1.11]
Kostandy 2008	5	0	-0.764 (0.719)		13.2 %	-0.76 [-2.17, 0.65]
Total (95% CI)					**100.0 %**	**-0.77 [-1.50, -0.03]**

Heterogeneity: $Tau^2 = 0.55$; $Chi^2 = 27.81$, df = 5 (P = 0.00004); $I^2 = 82\%$

Test for overall effect: Z = 2.05 (P = 0.041)

Test for subgroup differences: Not applicable



Analysis 3.1. Comparison 3 Neonate X kangaroo care X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 3 Neonate X kangaroo care X reactivity

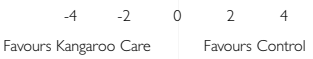
Outcome: 1 Pain-related distress reactivity

Study or subgroup	Kangaroo Care N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Chermont 2009	160	160	0.112 (0.112)		50.6 %	0.11 [-0.11, 0.33]
Kashaninia 2008	50	50	-1.923 (0.242)		49.4 %	-1.92 [-2.40, -1.45]
Total (95% CI)					**100.0 %**	**-0.89 [-2.89, 1.10]**

Heterogeneity: $Tau^2 = 2.04$; $Chi^2 = 58.24$, df = 1 (P<0.00001); $I^2 = 98\%$

Test for overall effect: Z = 0.88 (P = 0.38)

Test for subgroup differences: Not applicable



Analysis 4.1. Comparison 4 Neonate X kangaroo care X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 4 Neonate X kangaroo care X immediate regulation

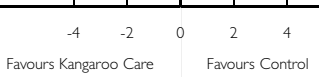
Outcome: 1 Pain-related distress regulation

Study or subgroup	Kangaroo Care N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Chermont 2009	160	160	-0.192 (0.112)		58.0 %	-0.19 [-0.41, 0.03]
Gray 2000	12	11	-1.304 (0.46)		42.0 %	-1.30 [-2.21, -0.40]
Total (95% CI)					**100.0 %**	**-0.66 [-1.73, 0.42]**

Heterogeneity: $Tau^2 = 0.51$; $Chi^2 = 5.52$, df = 1 (P = 0.02); $I^2 = 82\%$

Test for overall effect: Z = 1.20 (P = 0.23)

Test for subgroup differences: Not applicable



Analysis 5.1. Comparison 5 Preterm X swaddling/tucking-related X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 5 Preterm X swaddling/tucking-related X reactivity

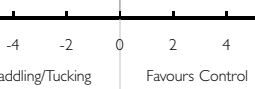
Outcome: 1 Pain-related distress reactivity

Study or subgroup	Swaddling/Tucking N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Axelin 2006	20	0	-1.51 (0.463)		12.9 %	-1.51 [-2.42, -0.60]
Axelin 2009	20	0	-0.952 (0.381)		14.0 %	-0.95 [-1.70, -0.21]
Comaru 2009	30	0	-0.913 (0.408)		13.6 %	-0.91 [-1.71, -0.11]
Comaru 2009	17	0	-0.551 (0.277)		15.3 %	-0.55 [-1.09, -0.01]
Hill 2005	12	0	-1.08 (0.363)		14.2 %	-1.08 [-1.79, -0.37]
Stevens 1999	122	0	0.138 (0.129)		16.6 %	0.14 [-0.11, 0.39]
Ward-Larson 2004	40	0	-2.231 (0.418)		13.5 %	-2.23 [-3.05, -1.41]
Total (95% CI)					**100.0 %**	**-0.97 [-1.63, -0.31]**

Heterogeneity: $Tau^2 = 0.67$; $Chi^2 = 49.75$, df = 6 (P<0.00001); $I^2 = 88\%$

Test for overall effect: Z = 2.87 (P = 0.0041)

Test for subgroup differences: Not applicable



Analysis 6.1. Comparison 6 Preterm X swaddling/tucking-related X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 6 Preterm X swaddling/tucking-related X immediate regulation

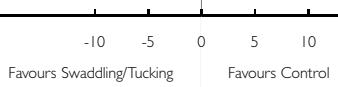
Outcome: 1 Pain-related distress regulation

Study or subgroup	Swaddling/Tucking N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Axelin 2006	20	0	-0.552 (0.337)		34.9 %	-0.55 [-1.21, 0.11]
Corff 1995	30	0	-0.943 (0.31)		41.3 %	-0.94 [-1.55, -0.34]
Fearon 1997	15	0	-0.707 (0.408)		23.8 %	-0.71 [-1.51, 0.09]
Total (95% CI)					**100.0 %**	**-0.75 [-1.14, -0.36]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.74$, df = 2 (P = 0.69); $I^2 = 0.0\%$

Test for overall effect: Z = 3.77 (P = 0.00016)

Test for subgroup differences: Not applicable



Analysis 7.1. Comparison 7 Neonate X swaddling/tucking X reactivity, Outcome 1 Pain-related distress.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 7 Neonate X swaddling/tucking X reactivity

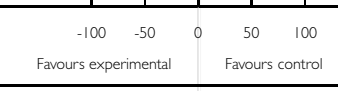
Outcome: 1 Pain-related distress

Study or subgroup	Swaddling/Tucking N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Morrow 2010	22	20	-1.26 (0.338)		100.0 %	-1.26 [-1.92, -0.60]
Total (95% CI)					**100.0 %**	**-1.26 [-1.92, -0.60]**

Heterogeneity: not applicable

Test for overall effect: Z = 3.73 (P = 0.00019)

Test for subgroup differences: Not applicable



Analysis 8.1. Comparison 8 Preterm X non-nutritive sucking-related X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 8 Preterm X non-nutritive sucking-related X reactivity

Outcome: 1 Pain-related distress reactivity

Study or subgroup	Sucking Related N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Bellieni 2001	17	0	-0.636 (0.376)		9.5 %	-0.64 [-1.37, 0.10]
Corbo 2000	26	0	-1.518 (0.407)		8.4 %	-1.52 [-2.32, -0.72]
Elserafy 2009	36	0	-0.236 (0.239)		17.1 %	-0.24 [-0.70, 0.23]
Elserafy 2009	0	0	-0.236 (0.238)		17.2 %	-0.24 [-0.70, 0.23]
Liaw 2010	52	52	-0.217 (0.197)		20.6 %	-0.22 [-0.60, 0.17]
Stevens 1999	122	0	-0.383 (0.133)		27.2 %	-0.38 [-0.64, -0.12]
Total (95% CI)					**100.0 %**	**-0.42 [-0.68, -0.15]**

Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 9.69$, df = 5 (P = 0.08); $I^2 = 48\%$

Test for overall effect: Z = 3.11 (P = 0.0019)

Test for subgroup differences: Not applicable

Analysis 9.1. Comparison 9 Preterm X non-nutritive sucking-related X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 9 Preterm X non-nutritive sucking-related X immediate regulation

Outcome: 1 Pain-related distress regulation

Study or subgroup	Sucking Related N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Corbo 2000	26	0	-0.784 (0.317)		11.9 %	-0.78 [-1.41, -0.16]
Elserafy 2009	0	0	-0.235 (0.239)		21.0 %	-0.24 [-0.70, 0.23]
Elserafy 2009	36	0	-0.229 (0.239)		21.0 %	-0.23 [-0.70, 0.24]
Liaw 2010	52	52	-0.374 (0.198)		30.6 %	-0.37 [-0.76, 0.01]
Whipple 2004	20	0	-0.202 (0.388)		8.0 %	-0.20 [-0.96, 0.56]
Whipple 2004	20	20	-0.776 (0.4)		7.5 %	-0.78 [-1.56, 0.01]
Total (95% CI)					**100.0 %**	**-0.38 [-0.59, -0.17]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.58$, df = 5 (P = 0.61); $I^2 = 0.0\%$

Test for overall effect: Z = 3.47 (P = 0.00053)

Test for subgroup differences: Not applicable

Analysis 10.1. Comparison 10 Neonate X non-nutritive sucking-related X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 10 Neonate X non-nutritive sucking-related X reactivity

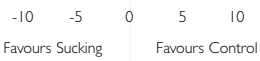
Outcome: 1 Pain-related distress reactivity

Study or subgroup	Sucking N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Bellieni 2002	20	20	-1.812 (0.375)		23.8 %	-1.81 [-2.55, -1.08]
Carbajal 1999	25	25	-1.963 (0.344)		24.5 %	-1.96 [-2.64, -1.29]
Liu 2010	35	35	-1.852 (0.286)		25.6 %	-1.85 [-2.41, -1.29]
Yilmaz 2010	30	30	-0.261 (0.259)		26.1 %	-0.26 [-0.77, 0.25]
Total (95% CI)					**100.0 %**	**-1.45 [-2.34, -0.57]**

Heterogeneity: $Tau^2 = 0.72$; $Chi^2 = 25.39$, df = 3 (P = 0.00001); $I^2 = 88\%$

Test for overall effect: Z = 3.21 (P = 0.0013)

Test for subgroup differences: Not applicable



Analysis 11.1. Comparison 11 Neonate X non-nutritive sucking-related X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 11 Neonate X non-nutritive sucking-related X immediate regulation

Outcome: 1 Pain-related distress regulation

Study or subgroup	Sucking N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Blass 1999	10	10	-0.904 (0.452)		14.1 %	-0.90 [-1.79, -0.02]
Bo 2000	27	0	-11.64 (2.257)		1.9 %	-11.64 [-16.06, -7.22]
Campos 1994	20	20	-0.97 (0.334)		16.0 %	-0.97 [-1.62, -0.32]
Greenberg 2002	21	21	-0.06 (0.309)		16.4 %	-0.06 [-0.67, 0.55]
Im 2008	33	33	-0.279 (0.247)		17.3 %	-0.28 [-0.76, 0.21]
Liu 2010	35	35	-1.354 (0.265)		17.1 %	-1.35 [-1.87, -0.83]
Yilmaz 2010	30	30	-0.6 (0.264)		17.1 %	-0.60 [-1.12, -0.08]
Total (95% CI)					**100.0 %**	**-0.90 [-1.54, -0.25]**

Heterogeneity: $Tau^2 = 0.56$; $Chi^2 = 37.78$, df = 6 ($P<0.00001$); $I^2 = 84\%$

Test for overall effect: Z = 2.73 (P = 0.0063)

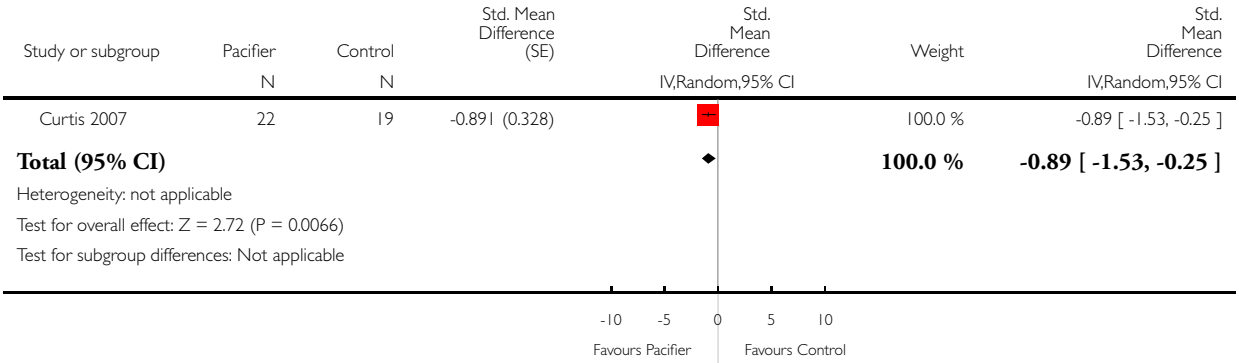
Test for subgroup differences: Not applicable

Analysis 12.1. Comparison 12 Older infants X non-nutritive sucking-related X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 12 Older infants X non-nutritive sucking-related X immediate regulation

Outcome: 1 Pain-related distress regulation

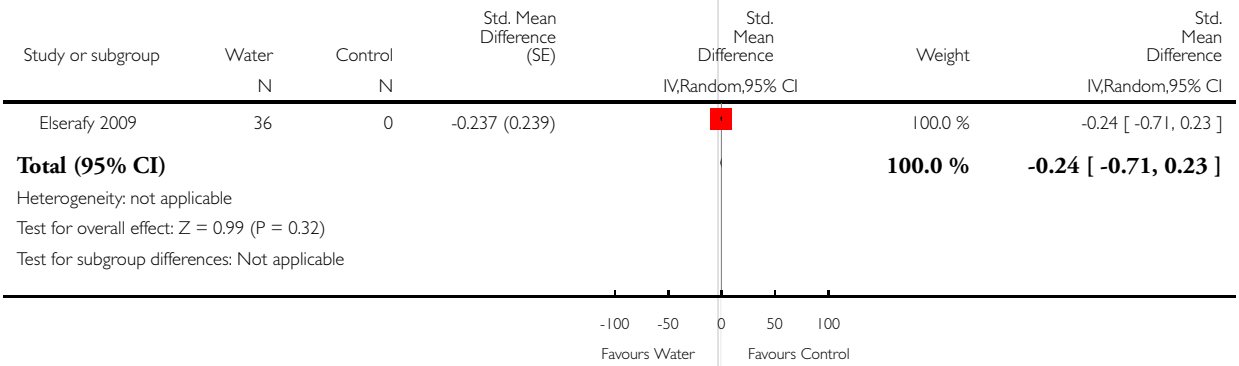


Analysis 13.1. Comparison 13 Preterm X swallowing water X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 13 Preterm X swallowing water X reactivity

Outcome: 1 Pain-related distress reactivity

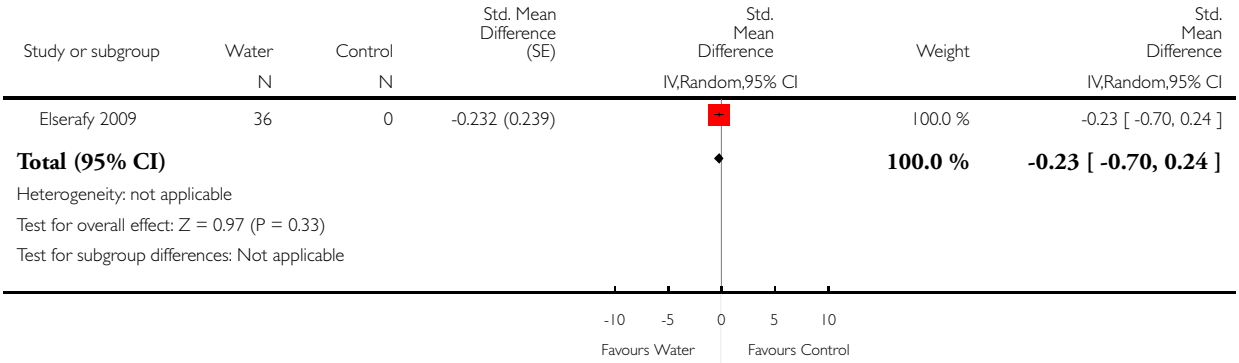


Analysis 14.1. Comparison 14 Preterm X swallowing water X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 14 Preterm X swallowing water X immediate regulation

Outcome: 1 Pain-related distress regulation

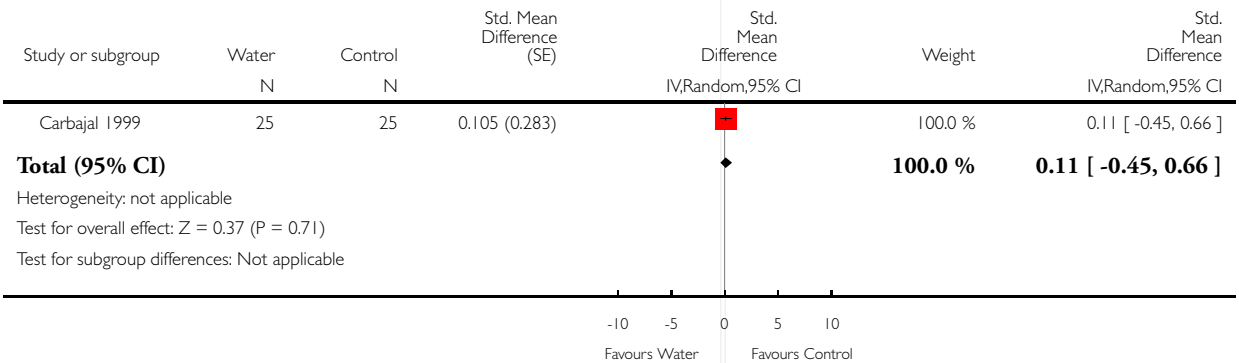


Analysis 15.1. Comparison 15 Neonate X swallowing water X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 15 Neonate X swallowing water X reactivity

Outcome: 1 Pain-related distress reactivity

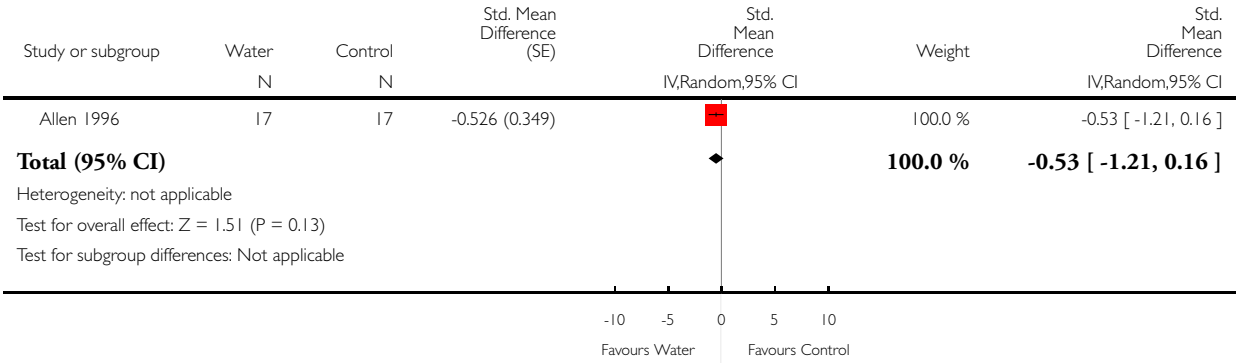


Analysis 16.1. Comparison 16 Neonate X swallowing water X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 16 Neonate X swallowing water X immediate regulation

Outcome: 1 Pain-related distress regulation

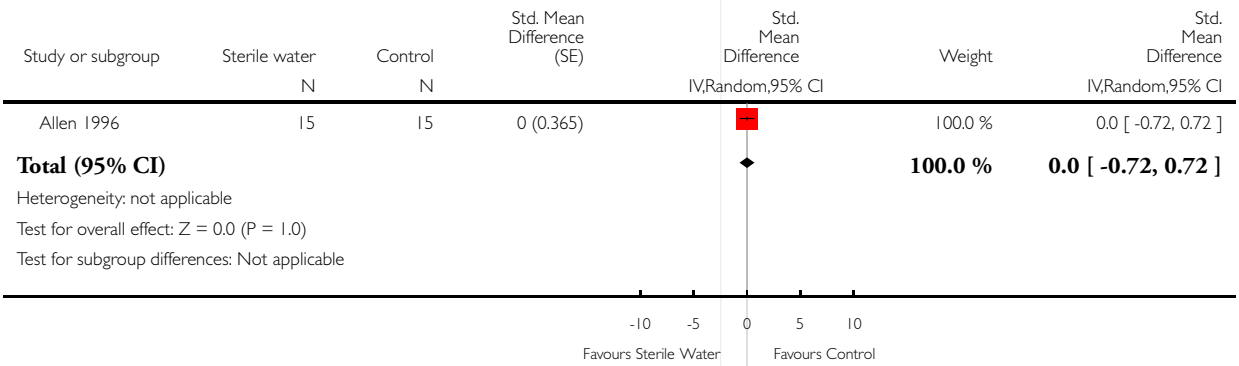


Analysis 17.1. Comparison 17 Older infants X swallowing water X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 17 Older infants X swallowing water X immediate regulation

Outcome: 1 Pain-related distress regulation

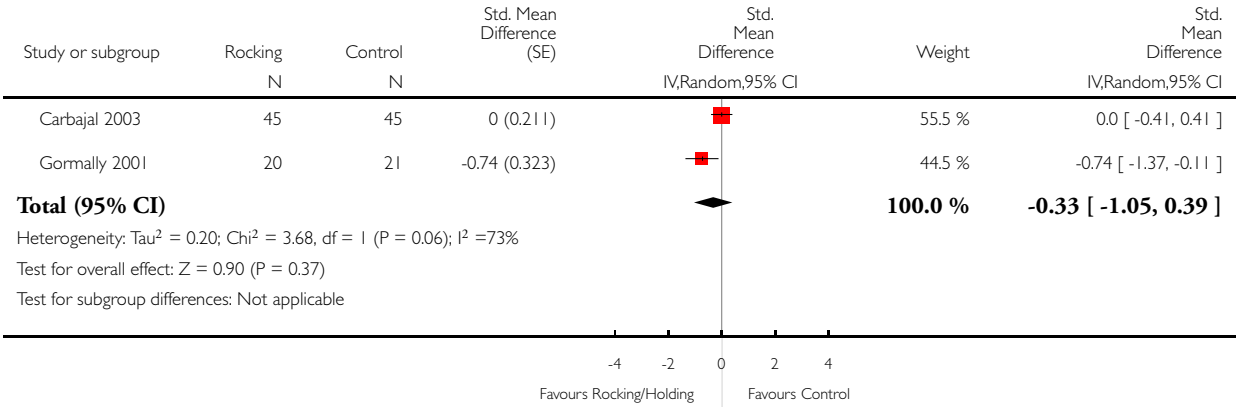


Analysis 18.1. Comparison 18 Neonate X rocking/holding X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 18 Neonate X rocking/holding X reactivity

Outcome: 1 Pain-related distress reactivity

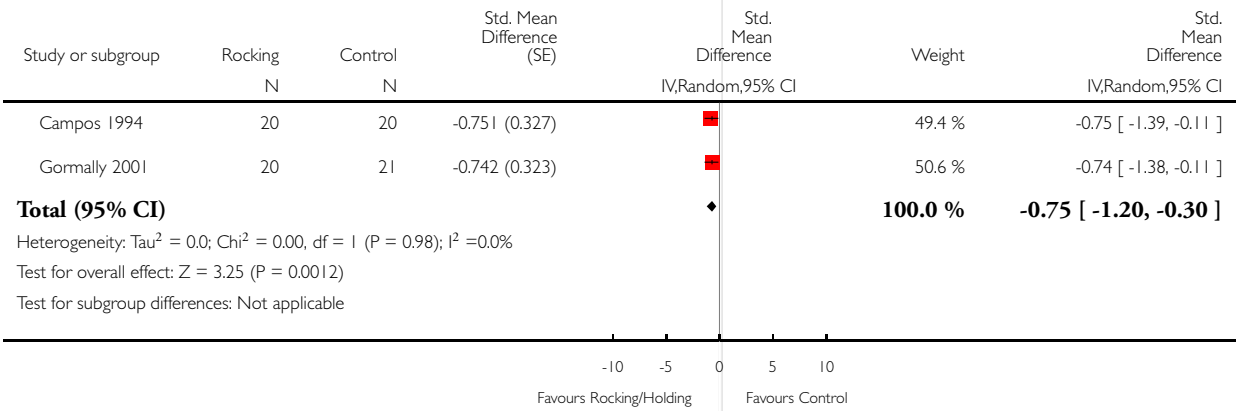


Analysis 19.1. Comparison 19 Neonate X rocking/holding X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 19 Neonate X rocking/holding X immediate regulation

Outcome: 1 Pain-related distress regulation



Analysis 20.1. Comparison 20 Older infants X rocking/holding X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 20 Older infants X rocking/holding X reactivity

Outcome: 1 Pain-related distress reactivity

Study or subgroup	Holding N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Ipp 2004	56	50	0.234 (0.195)		100.0 %	0.23 [-0.15, 0.62]
Total (95% CI)					**100.0 %**	**0.23 [-0.15, 0.62]**

Heterogeneity: not applicable

Test for overall effect: Z = 1.20 (P = 0.23)

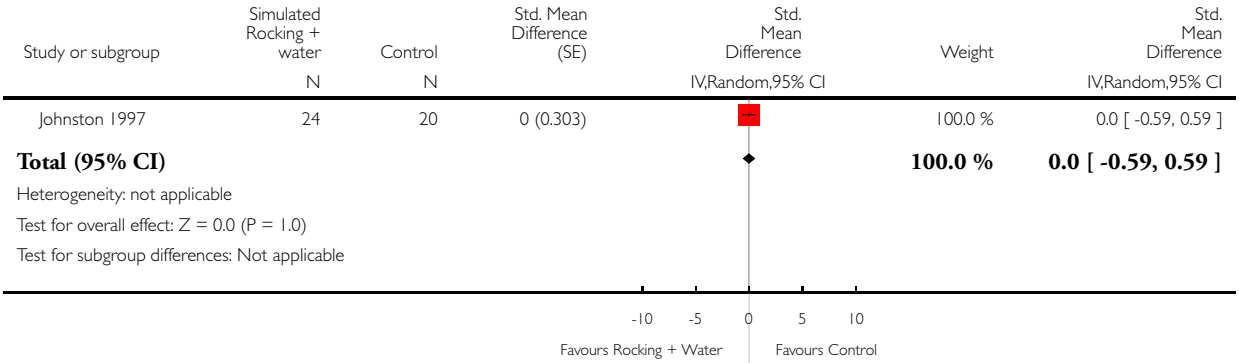
Test for subgroup differences: Not applicable

Analysis 21.1. Comparison 21 Preterm X simulated rocking + water X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 21 Preterm X simulated rocking + water X reactivity

Outcome: 1 Pain-related distress reactivity

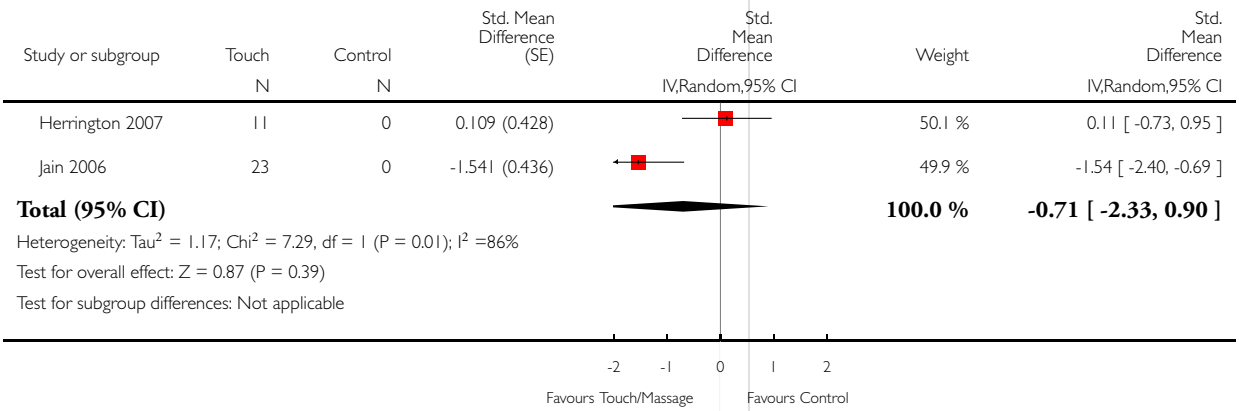


Analysis 22.1. Comparison 22 Preterm X touch or massage-related X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 22 Preterm X touch or massage-related X immediate regulation

Outcome: 1 Pain-related distress regulation

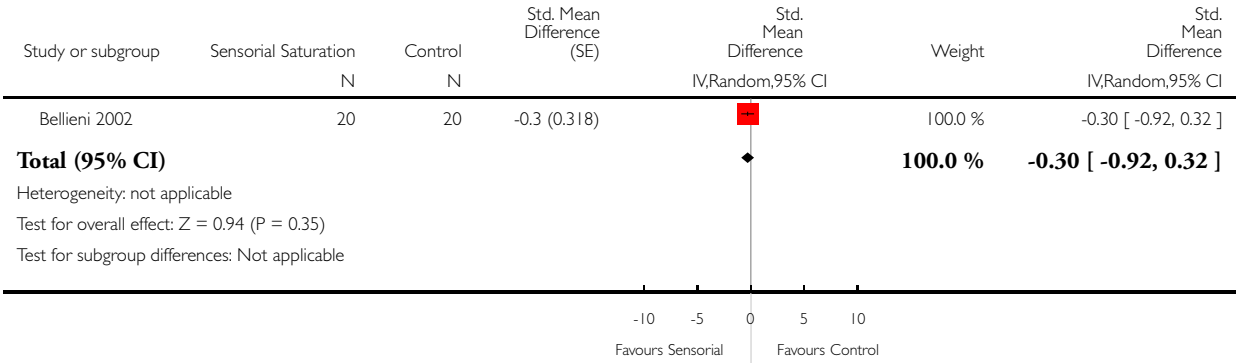


Analysis 23.1. Comparison 23 Neonate X touch or massage-related X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 23 Neonate X touch or massage-related X reactivity

Outcome: 1 Pain-related distress reactivity

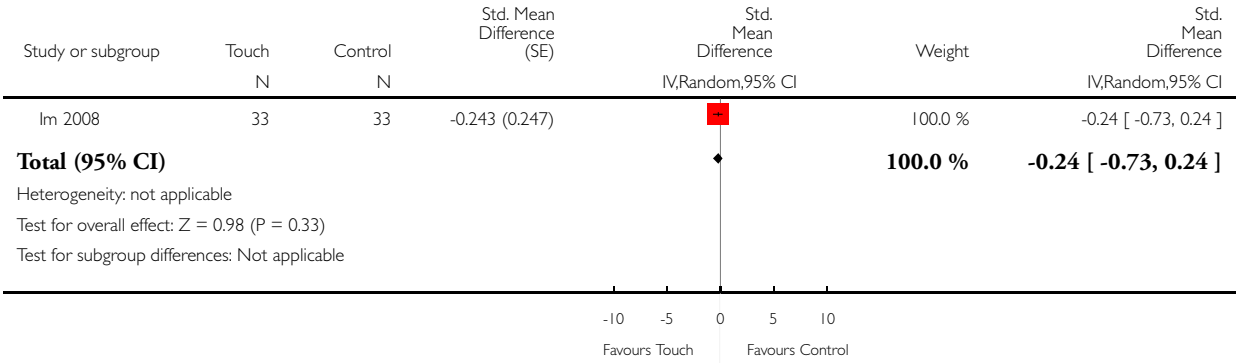


Analysis 24.1. Comparison 24 Neonate X touch or massage-related X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 24 Neonate X touch or massage-related X immediate regulation

Outcome: 1 Pain-related distress regulation

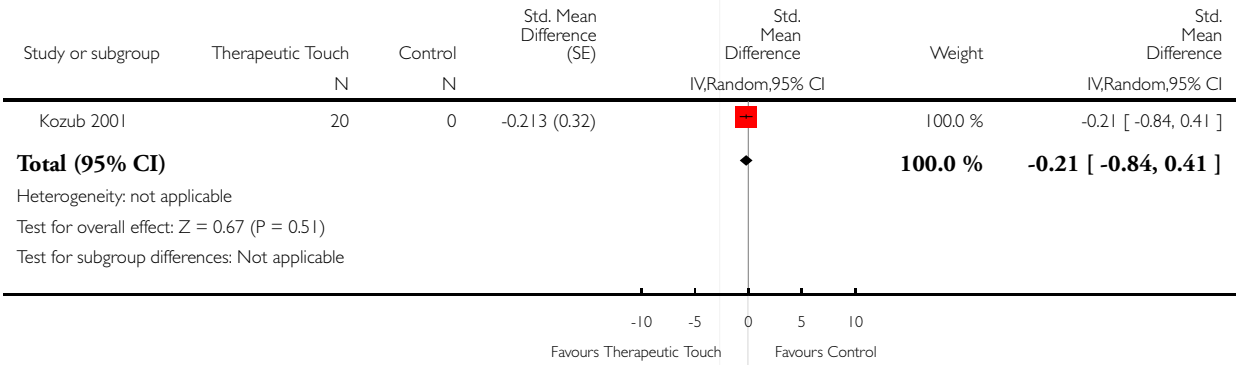


Analysis 25.1. Comparison 25 Older infants X touch or massage-related X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 25 Older infants X touch or massage-related X reactivity

Outcome: 1 Pain-related distress reactivity



Analysis 26.1. Comparison 26 Preterm X environment modification X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 26 Preterm X environment modification X reactivity

Outcome: 1 Pain-related distress reactivity

Study or subgroup	Environment Modification N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Catelin 2005	45	0	-12.037 (1.807)		48.7 %	-12.04 [-15.58, -8.50]
Sizun 2002	19	0	-1.12 (0.414)		51.3 %	-1.12 [-1.93, -0.31]
Total (95% CI)					**100.0 %**	**-6.44 [-17.13, 4.26]**

Heterogeneity: $Tau^2 = 57.87$; $Chi^2 = 34.68$, df = 1 (P<0.00001); I^2 =97%

Test for overall effect: Z = 1.18 (P = 0.24)

Test for subgroup differences: Not applicable

Analysis 27.1. Comparison 27 Preterm X environment modification X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 27 Preterm X environment modification X immediate regulation

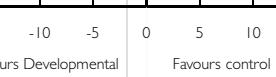
Outcome: 1 Pain-related distress regulation

Study or subgroup	Developmental Care N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Catelin 2005	45	0	-4.013 (0.634)		100.0 %	-4.01 [-5.26, -2.77]
Total (95% CI)					**100.0 %**	**-4.01 [-5.26, -2.77]**

Heterogeneity: not applicable

Test for overall effect: Z = 6.33 (P < 0.00001)

Test for subgroup differences: Not applicable



Analysis 28.1. Comparison 28 Older infants X toy distraction X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 28 Older infants X toy distraction X reactivity

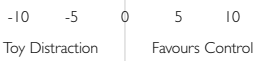
Outcome: 1 Pain-related distress reactivity

Study or subgroup	Toy Distraction N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Cramer-Berness 2005	40	41	-0.406 (0.225)		32.0 %	-0.41 [-0.85, 0.03]
Cramer-Berness 2005b	41	38	0.042 (0.225)		32.0 %	0.04 [-0.40, 0.48]
Hillgrove Stuart 2008	33	34	0.035 (0.3)		18.0 %	0.04 [-0.55, 0.62]
Hillgrove Stuart 2008	32	0	0.0304 (0.299)		18.1 %	0.03 [-0.56, 0.62]
Total (95% CI)					**100.0 %**	**-0.10 [-0.35, 0.14]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 2.64$, df = 3 (P = 0.45); $I^2 = 0.0\%$

Test for overall effect: Z = 0.82 (P = 0.41)

Test for subgroup differences: Not applicable



Analysis 29.1. Comparison 29 Older infants X toy distraction X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 29 Older infants X toy distraction X immediate regulation

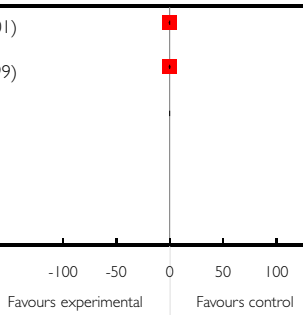
Outcome: 1 Pain-related distress regulation

Study or subgroup	Experimental N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Hillgrove Stuart 2008	33	34	-0.167 (0.301)		49.7 %	-0.17 [-0.76, 0.42]
Hillgrove Stuart 2008	32	34	0.004 (0.299)		50.3 %	0.00 [-0.58, 0.59]
Total (95% CI)					**100.0 %**	**-0.08 [-0.50, 0.33]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.16$, df = 1 (P = 0.69); $I^2 = 0.0\%$

Test for overall effect: Z = 0.38 (P = 0.70)

Test for subgroup differences: Not applicable



Analysis 30.1. Comparison 30 Older infants X video distraction X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 30 Older infants X video distraction X reactivity

Outcome: 1 Pain-related distress reactivity

Study or subgroup	Video/Toy Distraction N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Cohen 2002	49	41	-0.701 (0.218)		100.0 %	-0.70 [-1.13, -0.27]
Total (95% CI)					**100.0 %**	**-0.70 [-1.13, -0.27]**

Heterogeneity: not applicable

Test for overall effect: Z = 3.22 (P = 0.0013)

Test for subgroup differences: Not applicable

Analysis 31.1. Comparison 31 Older infants X video distraction X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 31 Older infants X video distraction X immediate regulation

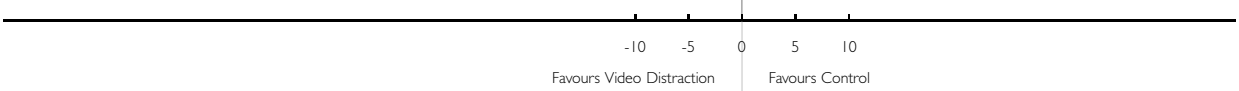
Outcome: 1 Pain-related distress regulation

Study or subgroup	Video distraction N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Cohen 2006	63	63	-0.836 (0.186)		100.0 %	-0.84 [-1.20, -0.47]
Total (95% CI)					**100.0 %**	**-0.84 [-1.20, -0.47]**

Heterogeneity: not applicable

Test for overall effect: Z = 4.49 (P < 0.00001)

Test for subgroup differences: Not applicable



Analysis 32.1. Comparison 32 Older infant X structured parent involvement X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 32 Older infant X structured parent involvement X reactivity

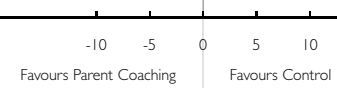
Outcome: 1 Pain-related distress reactivity

Study or subgroup	Parent Coaching N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Bustos 2008	25	25	-0.313 (0.285)		29.2 %	-0.31 [-0.87, 0.25]
Cramer-Berness 2005	42	41	-0.596 (0.224)		35.8 %	-0.60 [-1.04, -0.16]
Cramer-Berness 2005b	38	38	0.115 (0.23)		35.1 %	0.12 [-0.34, 0.57]
Total (95% CI)					**100.0 %**	**-0.26 [-0.70, 0.17]**

Heterogeneity: $Tau^2 = 0.09$; $Chi^2 = 4.94$, df = 2 (P = 0.08); $I^2 = 60\%$

Test for overall effect: Z = 1.19 (P = 0.23)

Test for subgroup differences: Not applicable



Analysis 33.1. Comparison 33 Older infants X structured parent involvement X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 33 Older infants X structured parent involvement X immediate regulation

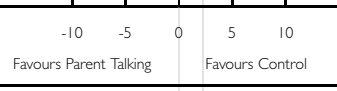
Outcome: 1 Pain-related distress regulation

Study or subgroup	Parent Talking Related N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Bauchner 1996	157	131	0.019 (0.118)		100.0 %	0.02 [-0.21, 0.25]
Total (95% CI)					**100.0 %**	**0.02 [-0.21, 0.25]**

Heterogeneity: not applicable

Test for overall effect: Z = 0.16 (P = 0.87)

Test for subgroup differences: Not applicable



Analysis 34.1. Comparison 34 Preterm X mother's voice X reactivity, Outcome 1 Pain-related distress reactivity.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 34 Preterm X mother's voice X reactivity

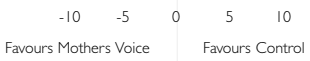
Outcome: 1 Pain-related distress reactivity

Study or subgroup	Mothers Voice N	Control N	Std. Mean Difference (SE)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Johnston 2007	19	0	-0.294 (0.331)		100.0 %	-0.29 [-0.94, 0.35]
Total (95% CI)					**100.0 %**	**-0.29 [-0.94, 0.35]**

Heterogeneity: not applicable

Test for overall effect: Z = 0.89 (P = 0.37)

Test for subgroup differences: Not applicable

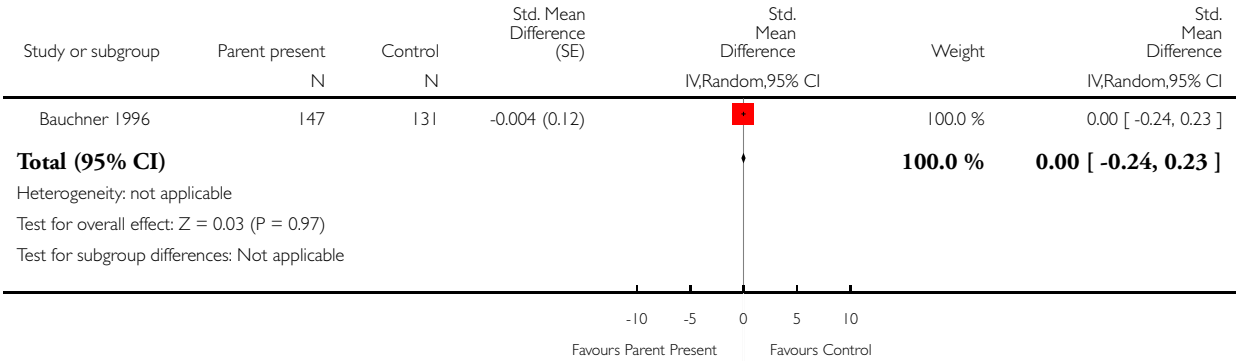


Analysis 35.1. Comparison 35 Older infants X parent present X immediate regulation, Outcome 1 Pain-related distress regulation.

Review: Non-pharmacological management of infant and young child procedural pain

Comparison: 35 Older infants X parent present X immediate regulation

Outcome: 1 Pain-related distress regulation



ADDITIONAL TABLES

Table 1. Procedures analyzed under scope of current review

Name	Description
Heel-stick	Heel-stick involves lancing of the lateral aspect of the infant's heel, squeezing the heel, and collecting the pooled blood
Venipuncture (also known as venepuncture)	Surgical puncture of a vein, especially for the withdrawal of blood or for intravenous medication
Needle	The act of forcing a liquid (such as vaccines) into tissue using a sharp needle
Diaper change (preterm only)	Care-giving intervention in which soiled diapers are changed and causes distress that some suspect to be pain-related
Endotracheal suctioning	A component of bronchial hygiene for mechanical ventilation and involves the mechanical aspiration of pulmonary secretion from the intubated airway. Its primary purpose is to remove airway secretions in order to prevent obstructions
Weighing procedure (preterm only)	Routine care procedure in which the preterm is weighed on a scale and causes distress that some suspect to be pain-related

Table 2. Summary of overall findings

	Preterm Infants		**Neonates**		**Older Infants**	
	Reactivity	Immediate regulation	Reactivity	Immediate regulation	Reactivity	Immediate regulation
Kangaroo care	**1**	**1**	3	3	-	-
Non-nutritive sucking-related	**1**	**1**	**1**	**1**	-	2
Swaddling/ tucking-related	**1**	**1**	2	-	-	-
Touch or massage-related	-	3	3	3	3	-
Environment modification	3	2	-	-	-	-
Simulated rocking and water	3	-	-	-	-	-
Simulated Mother's voice	3	-	-	-	-	-
Swallowing water	3	3	3	3	-	3
Rocking or holding	-	-	3	**1**	3	-
Toy distraction	-	-	-	-	4	3
Video distraction	-	-	-	-	2	2
Parent present	-	-	-	-	-	3
Structured parent involvement	-	-	-	-	3	3

Legend:

1. Sufficient evidence supports efficacy for reducing pain-related behaviors (support of two or more trials).
2. Limited evidence suggests efficacy for reducing pain-related behaviors (e.g. support of 1 trial or heterogeneity among trials).
3. Limited evidence suggests inefficacy for reducing pain-related behaviors (e.g. support of 1 trial or heterogeneity among trials).
4. Sufficient evidence supports inefficacy for reducing pain-related behaviors (support of two or more trials).

Table 3. Kangaroo care - pain-related distress

Age	Pain responses	No. of treatment arms	No. of participants	Statistical method	Heterogeneity	Effect size
Preterm	Reactivity	5	236	Std. Mean Difference (IV, Random, 95% CI)	P < 0.00001, sig.*	-1.12 (-2.04 to -0.21)
	Immediate regulation	5	222	Std. Mean Difference (IV, Random, 95% CI)	P < 0.0001, sig.*	-0.77 (-1.50 to -0.03)
Neonate	Reactivity	2	420	Std. Mean Difference (IV, Random, 95% CI)	P < 0.00001, sig.*	-0.89 (-2.89 to 1.10)
	Immediate regulation	2	343	Std. Mean Difference (IV, Random, 95% CI)	P = 0.02, sig*	-0.66 (-1.73 to 0.42)
Older Infant	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---

Table 4. Swaddling/tucking - pain-related distress

Age	Pain responses	No. of treatment arms	No. of participants	Statistical method	Heterogeneity	Effect size
Preterm	Reactivity	6	261	Std. Mean Difference (IV, Random, 95% CI)	P < 0.00001, sig*	-0.97 (-1.63 to -0.31)
	Immediate regulation	3	65	Std. Mean Difference (IV, Random, 95% CI)	P = 0.69	-0.75 (-1.14 to -0.36)
Neonate	Reactivity	1	42	Std. Mean Difference (IV, Random, 95% CI)	n/a	-1.26 (-1.92 to -0.60)
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	---	---	---	---	---

Table 4. Swaddling/tucking - pain-related distress (*Continued*)

	Immediate regulation	---	---	---	---	---

Table 5. Non-nutritive sucking-related - pain-related distress

Age	Pain responses	No. of treatment arms	No. of participants	Statistical method	Heterogeneity	Effect size
Preterm	Reactivity	6	305	Std. Mean Difference (IV, Random, 95% CI)	P = 0.08	-0.42 (-0.68 to -0.15)
	Immediate regulation	6	226	Std. Mean Difference (IV, Random, 95% CI)	P = 0.61	-0.38 (-0.59 to -0.17)
Neonate	Reactivity	4	220	Std. Mean Difference (IV, Random, 95% CI)	P < 0.00001, sig*	-1.45 (-2.34 to -0.57)
	Immediate regulation	7	325	Std. Mean Difference (IV, Random, 95% CI)	P < 0.00001, sig*	-0.90 (-1.54 to -0.25)
Older Infant	Reactivity	---	---	---	---	---
	Immediate regulation	1	41	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.89 (-1.53 to -0.25)

Table 6. Swallowing water - pain-related distress

Age	Pain responses	No. of treatment arms	No. of participants	Statistical method	Heterogeneity	Effect size
Preterm	Reactivity	1	36	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.24 (-0.71 to 0.23)
	Immediate regulation	1	36	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.23 (-0.70 to 0.24)
Neonate	Reactivity	1	50	Std. Mean Difference (IV, Random, 95% CI)	n/a	0.10 (-0.45 to 0.66)

Table 6. Swallowing water - pain-related distress (*Continued*)

	Immediate regulation	1	34	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.53 (-1.21 to 0.16)
Older Infant	Reactivity	---	---	---	---	---
	Immediate regulation	1	30	Std. Mean Difference (IV, Random, 95% CI)	n/a	0.00 (-0.72 to 0.72)

Table 7. Rocking/holding - pain-related distress

Age	**Pain responses**	**No. of treatment arms**	No. of participants	**Statistical method**	**Heterogeneity**	**Effect size**
Preterm	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Neonate	Reactivity	2	131	Std. Mean Difference (IV, Random, 95% CI)	P = 0.06	-0.33 (-1.05 to 0.39)
	Immediate regulation	2	81	Std. Mean Difference (IV, Random, 95% CI)	P = 0.98	-0.75 (-1.20 to -0.30)
Older Infant	Reactivity	1	106	Std. Mean Difference (IV, Random, 95% CI)	n/a	0.23 (-0.15 to 0.62)
	Immediate regulation	---	---	---	---	---

Table 8. Simulated rocking + water - pain-related distress

Age	**Pain responses**	**No. of treatment arms**	No. of participants	**Statistical method**	**Heterogeneity**	**Effect size**
Preterm	Reactivity	1	44	Std. Mean Difference (IV, Random, 95% CI)	n/a	0.00 (-0.59 to 0.59)
	Immediate regulation	---	---	---	---	---

Table 8. Simulated rocking + water - pain-related distress (*Continued*)

Neonate	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---

Table 9. Touch or massage-related - pain-related distress

Age	Pain responses	No. of treatment arms	No. of participants	Statistical method	Heterogeneity	Effect size
Preterm	Reactivity	---	---	---	---	---
	Immediate regulation	2	34	Std. Mean Difference (IV, Random, 95% CI)	P = 0.007, sig*	-0.71 (-2.33 to 0.90)
Neonate	Reactivity	1	40	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.30 (-0.92 to 0.32)
	Immediate regulation	1	66	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.24 (-0.73 to 0.24)
Older Infant	Reactivity	1	20	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.21 (-0.84 to 0.41)
	Immediate regulation	---	---	---	---	---

Table 10. Environment modification - pain-related distress

Age	Pain responses	No. of treatment arms	No. of participants	Statistical method	Heterogeneity	Effect size
Preterm	Reactivity	2	64	Std. Mean Difference (IV, Random, 95% CI)	P < 0.00001, sig*	-6.44 (-17.13 to 4.26)

Table 10. Environment modification - pain-related distress (*Continued*)

	Immediate regulation	1	45	Std. Mean Difference (IV, Random, 95% CI)	n/a	-4.01 (-5.26 to -2.77)
Neonate	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---

Table 11. Toy distraction - pain-related distress

Age	**Pain responses**	**No. of treatment arms**	No. of participants	**Statistical method**	**Heterogeneity**	**Effect size**
Preterm	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Neonate	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	4	259	Std. Mean Difference (IV, Random, 95% CI)	P = 0.45	-0.10 (-0.35 to 0.14)
	Immediate regulation	2	99	Std. Mean Difference (IV, Random, 95% CI)	P = 0.69	-0.08 (-0.50 to 0.33)

Table 12. Video distraction - pain-related distress

Age	**Pain responses**	**No. of treatment arms**	No. of participants	**Statistical method**	**Heterogeneity**	**Effect size**
Preterm	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---

Table 12. Video distraction - pain-related distress (*Continued*)

Neonate	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	1	90	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.70 (-1.13 to -0.27)
	Immediate regulation	1	126	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.84 (-1.20 to -0.47)

Table 13. Structured parent involvement - pain-related distress

Age	**Pain responses**	**No. of treatment arms**	No. of participants	**Statistical method**	**Heterogeneity**	**Effect size**
Preterm	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Neonate	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	3	209	Std. Mean Difference (IV, Random, 95% CI)	P = 0.08	-0.26 (-0.70 to 0.17)
	Immediate regulation	1	288	Std. Mean Difference (IV, Random, 95% CI)	n/a	0.02 (-0.21 to 0.25)

Table 14. Mothers voice - pain-related distress

Age	**Pain responses**	**No. of treatment arms**	No. of participants	**Statistical method**	**Heterogeneity**	**Effect size**
Preterm	Reactivity	1	19	Std. Mean Difference (IV, Random, 95% CI)	n/a	-0.29 (-0.94 to 0.35)
	Immediate regulation	---	---	---	---	---

Table 14. Mothers voice - pain-related distress (*Continued*)

Neonate	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---

Table 15. Parent present - pain-related distress

Age	**Pain responses**	**No. of treatment arms**	No. of participants	**Statistical method**	**Heterogeneity**	**Effect size**
Preterm	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Neonate	Reactivity	---	---	---	---	---
	Immediate regulation	---	---	---	---	---
Older Infant	Reactivity	---	---	---	---	---
	Immediate regulation	1	278	Std. Mean Difference (IV, Random, 95% CI)	n/a	0.00 (-0.24 to 0.23)

APPENDICES

Appendix 1. MEDLINE search strategy

This search strategy was adapted or special filters were applied for other databases. Other related key words and MeSH terms were included as appropriate depending on the terms used in each of the specific databases.

1. Needles/
2. needle*.mp.
3. (blood sampl* or immuni* or inoculat* or vaccin* or inject* or "finger prick*" or finger-prick or "heel prick*" or heel-prick* or "heel lance*" or heel-lance* or "heel puncture*" or heel-puncture* or "heel stick" or suture* or (laceration* adj3 repair*)).mp.
4. ("lumbar puncture" or lumbar-puncture* or "spinal tap*" or spinal-tap*).mp. subject heading word]
5. ("bone marrow aspiration" or "bone marrow biops*").mp.
6. (intravenous or intra-venous or venepuncture* or venipuncture* or venous cannulation* or (arterial blood gas* and cannul*)).mp.
7. ((catheter adj6 insert*) or catheter* or port-a-cath* or portacath).mp. subject heading word]
8. ("central line" adj6 (insert* or remov*)).mp.
9. (central venous catheter* adj6 insert*).mp.
10. (local analges* or local anaesthe* or local anesthe*).mp.
11. ((arterial puncture or artery) adj6 puncture*).mp.
12. "arterial line*".mp.
13. (thoracocentesis or paracentesis).mp.
14. or/1-13
15. exp Pain/
16. Pain Measurement/
17. PAIN THRESHOLD/
18. pain*.mp.
19. or/15-18
20. 14 and 19
21. ((vaccin* adj6 pain) or (cannul* adj6 pain) or (needle* adj6 pain*) or (needle* adj6 distress*) or (needle* adj6 discomfort) or (needle* adj6 fear*) or (needle* adj6 fright*) or (needle* adj6 anxious) or (needle* adj6 anxiet*) or (procedure* adj6 pain*) or (intervention* adj6 pain*) or (intervention* adj6 distress*) or (procedure adj6 distress*) or (procedure* adj6 discomfort*) or (procedure-related adj6 pain)).mp.
22. or/20-21
23. Pain, Postoperative/
24. ((postoperative adj3 pain*) or (post-operative adj3 pain*) or post-operative-pain).mp. subject heading word]
25. ((post-surgical adj3 pain*) or ("post surgical" adj3 pain*) or (post-surgery adj3 pain*) or (post adj surg* adj3 pain*)).mp.
26. (post* adj4 pain*).mp.
27. "pain relief after".mp.
28. ("pain following" adj3 surg*).mp.
29. (posttreatment adj3 pain*).mp.
30. ("pain control after" adj4 surg*).mp.
31. ((post surg* or post-surg*) and (pain* or discomfort)).mp.
32. ((pain* adj3 "after surg*") or (pain* adj3 "after operat*")).mp.
33. ((pain* adj3 "follow* operat*") or (pain* adj3 "follow* surg*")).mp. subject heading word]
34. or/23-33
35. 22 or 34
36. Child, Preschool/
37. exp Infant/
38. (baby or babies or neonate* or newborn or child* or infant* or paediatric* or pediatric*).mp.
39. or/36-38
40. 35 and 39

MEDLINE FILTER

Cochrane Highly Sensitive Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2009 revision); OVID format

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. drug therapy.fs.
6. randomly.ab.
7. trial.ab.
8. groups.ab.
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. humans.sh.
11. 9 and 10

Appendix 2. PsycINFO search strategy

1 Needles/
2 needle$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
3 (blood sampl* or immuni$ or inoculat$ or vaccin$ or inject$ or "finger prick$" or finger-prick or "heel prick$" or heel-prick$ or "heel lance$" or heel-lance$ or "heel puncture$" or heel-puncture$ or "heel stick" or suture$ or (laceration$ adj3 repair$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
4 ("lumbar puncture" or lumbar-puncture$ or "spinal tap$" or spinal-tap$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
5 ("bone marrow aspiration" or "bone marrow biops$").mp. [mp=title, abstract, heading word, table of contents, key concepts]
6 (intravenous or intra-venous or venepuncture$ or venipuncture$ or venous cannulation$ or (arterial blood gas$ and cannul$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
7 ((catheter adj6 insert$) or catheter$ or port-a-cath$ or portacath).mp. [mp=title, abstract, heading word, table of contents, key concepts]
8 ("central line" adj6 (insert$ or remov$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
9 (central venous catheter$ adj6 insert$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
10 (local analges$ or local anaesthe$ or local anesthe$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
11 ((arterial puncture or artery) adj6 puncture$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
12 "arterial line$".mp. [mp=title, abstract, heading word, table of contents, key concepts]
13 (thoracocentesis or paracentesis).mp. [mp=title, abstract, heading word, table of contents, key concepts]
14 or/1-13
15 exp Pain/
16 Pain Measurement/
17 PAIN THRESHOLD/
18 pain$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
19 or/15-18
20 14 and 19
21 ((vaccin* adj6 pain) or (cannul* adj6 pain) or (needle$ adj6 pain$) or (needle$ adj6 distress$) or (needle$ adj6 discomfort) or (needle$ adj6 fear$) or (needle$ adj6 fright$) or (needle$ adj6 anxious) or (needle$ adj6 anxiet$) or (procedure$ adj6 pain$) or (intervention$ adj6 pain$) or (intervention$ adj6 distress$) or (procedure adj6 distress$) or (procedure$ adj6 discomfort$) or (procedure-related adj6 pain)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
22 or/20-21
23 Pain, Postoperative/
24 ((postoperative adj3 pain$) or (post-operative adj3 pain$) or post-operative-pain).mp. [mp=title, abstract, heading word, table of contents, key concepts]
25 ((post-surgical adj3 pain$) or ("post surgical" adj3 pain$) or (post-surgery adj3 pain$) or (post adj surg$ adj3 pain$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
26 (post$ adj4 pain$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
27 "pain relief after".mp. [mp=title, abstract, heading word, table of contents, key concepts]
28 ("pain following" adj3 surg$).mp. [mp=title, abstract, heading word, table of contents, key concepts]

29 (posttreatment adj3 pain$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
30 ("pain control after" adj4 surg$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
31 ((post surg$ or post-surg$) and (pain$ or discomfort)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
32 ((pain$ adj3 "after surg$") or (pain$ adj3 "after operat$")).mp. [mp=title, abstract, heading word, table of contents, key concepts]
33 ((pain$ adj3 "follow$ operat$") or (pain$ adj3 "follow$ surg$")).mp. [mp=title, abstract, heading word, table of contents, key concepts]
34 or/23-33
35 22 or 34
36 Child, Preschool/
37 exp Infant/
38 (baby or babies or neonate$ or newborn or child$ or infant$ or paediatric$ or pediatric$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
39 or/36-38
40 35 and 39
41 limit 40 to yr="2009"
42 from 41 keep 1-3

Appendix 3. EMBASE search strategy

1 random*.ti,ab.
2 factorial*.ti,ab.
3 (crossover* or cross over* or cross-over*).ti,ab.
4 placebo*.ti,ab.
5 (doubl* adj blind*).ti,ab.
6 (singl* adj blind*).ti,ab.
7 assign*.ti,ab.
8 allocat*.ti,ab.
9 volunteer*.ti,ab.
10 CROSSOVER PROCEDURE.sh.
11 DOUBLE-BLIND PROCEDURE.sh.
12 RANDOMIZED CONTROLLED TRIAL.sh.
13 SINGLE BLIND PROCEDURE.sh.
14 or/1-13
15 ANIMAL/ or NONHUMAN/ or ANIMAL EXPERIMENT/
16 HUMAN/
17 15 and 16
18 15 not 17
19 14 not 18
20 Needles/
21 needle$.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
22 (blood sampl* or immuni$ or inoculat$ or vaccin$ or inject$ or "finger prick$" or finger-prick or "heel prick$" or heel-prick$ or "heel lance$" or heel-lance$ or "heel puncture$" or heel-puncture$ or "heel stick" or suture$ or (laceration$ adj3 repair$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
23 ("lumbar puncture" or lumbar-puncture$ or "spinal tap$" or spinal-tap$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
24 ("bone marrow aspiration" or "bone marrow biops$").mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
25 (intravenous or intra-venous or venepuncture$ or venipuncture$ or venous cannulation$ or (arterial blood gas$ and cannul$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
26 ((catheter adj6 insert$) or catheter$ or port-a-cath$ or portacath).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]

27 ("central line" adj6 (insert$ or remov$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
28 (central venous catheter$ adj6 insert$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
29 (local analges$ or local anaesthe$ or local anesthe$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
30 ((arterial puncture or artery) adj6 puncture$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
31 "arterial line$".mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
32 (thoracocentesis or paracentesis).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
33 or/20-32
34 exp Pain/
35 Pain Measurement/
36 PAIN THRESHOLD/
37 pain$.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
38 or/34-37
39 33 and 38
40 ((vaccin* adj6 pain) or (cannul* adj6 pain) or (needle$ adj6 pain$) or (needle$ adj6 distress$) or (needle$ adj6 discomfort) or (needle$ adj6 fear$) or (needle$ adj6 fright$) or (needle$ adj6 anxious) or (needle$ adj6 anxiet$) or (procedure$ adj6 pain$) or (intervention$ adj6 pain$) or (intervention$ adj6 distress$) or (procedure adj6 distress$) or (procedure$ adj6 discomfort$) or (procedure-related adj6 pain)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
41 or/39-40
42 Pain, Postoperative/
43 ((postoperative adj3 pain$) or (post-operative adj3 pain$) or post-operative-pain).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
44 ((post-surgical adj3 pain$) or ("post surgical" adj3 pain$) or (post-surgery adj3 pain$) or (post adj surg$ adj3 pain$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
45 (post$ adj4 pain$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
46 "pain relief after".mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
47 ("pain following" adj3 surg$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
48 (posttreatment adj3 pain$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
49 ("pain control after" adj4 surg$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
50 ((post surg$ or post-surg$) and (pain$ or discomfort)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
51 ((pain$ adj3 "after surg$") or (pain$ adj3 "after operat$")).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
52 ((pain$ adj3 "follow$ operat$") or (pain$ adj3 "follow$ surg$")).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
53 or/42-52
54 41 or 53
55 Child, Preschool/
56 exp Infant/
57 (baby or babies or neonate$ or newborn or child$ or infant$ or paediatric$ or pediatric$).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]

58 or/55-57
59 54 and 58
60 59 and 19
61 limit 60 to yr="2009"
62 from 61 keep 1-40

Filter for EMBASE (Ovid format) 2008

1. random*.ti,ab.
2. factorial*.ti,ab.
3. (crossover* or cross over* or cross-over*).ti,ab.
4. placebo*.ti,ab.
5. (doubl* adj blind*).ti,ab.
6. (singl* adj blind*).ti,ab.
7. assign*.ti,ab.
8. allocat*.ti,ab.
9. volunteer*.ti,ab.
10. CROSSOVER PROCEDURE.sh.
11. DOUBLE-BLIND PROCEDURE.sh.
12. RANDOMIZED CONTROLLED TRIAL.sh.
13. SINGLE BLIND PROCEDURE.sh.
14. or/1-13
15. ANIMAL/ or NONHUMAN/ or ANIMAL EXPERIMENT/
16. HUMAN/
17. 15 and 16
18. 15 not 17
19. 14 not 18

Appendix 4. CINAHL search strategy

S35. S31 and S34
S34. S32 or S33
S33. baby or babies or neonate* or newborn or child* or infant* or paediatric* or pediatric*
S32. (MH "Infant+") or (MH "Child, Preschool")
S31. S18 or S30
S30. S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29
S29. (pain* n3 "follow* operat*") or (pain* n3 "follow* surg*")
S28. (pain* n3 "after surg*") or (pain* n3 "after operat*")
S27. ("post surg*" or post-surg*) and (pain* or discomfort)
S26. "pain control after" n4 surg*
S25. posttreatment n3 pain*
S24. "pain following" n3 surg*
S23. "pain relief after"
S22. post* n4 pain*
S21. (post-surgical n3 pain*) or ("post surgical" n3 pain*) or (post-surgery n3 pain*) or (post n1 surg* n3 pain*)
S20. (postoperative n3 pain*) or (post-operative n3 pain*) or post-operative-pain
S19. (MH "Postoperative Pain")
S18. S16 or S17
S17. (needle* n6 pain*) or (needle* n6 distress*) or (needle* n6 discomfort) or (needle* n6 fear*) or (needle* n6 fright*) or (needle* n6 anxious) or (needle* n6 anxiet*) or (procedur* n6 pain*) or (intervention* n6 pain*) or (intervention* n6 distress*) or (procedur* n6 distress*) or (procedur* n6 discomfort*) or (procedure-related n6 pain*)
S16. S14 and S15
S15. (MH "Pain+") or (MH "Pain Measurement") or (MH "Pain Threshold") or pain*
S14. S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11

S12. or S13
S13. thoracocentesis or paracentesis
S12. "arterial line*"
S11. punctur* n6 arter*
S10. "local analges*" or "local anaesthe*" or "local anesthe*" modes
S9. "central venous catheter" n6 insert*
S8. "central line" n6 (insert* or remov*)
S7. (catheter n6 insert*) or catheter* or port-a-cath* or portacath
S6. intravenous or intra-venous or venepuncture* or venipuncture* or "venous cannulation*" or ("arterial blood gas*" and cannul*)
S5. "bone marrow aspiration" or "bone marrow biops*"
S4. "lumbar puncture" or lumbar-puncture* or "spinal tap*" or spinal-tap*
S3. (immuni* or inoculat* or vaccin* or inject* or "finger prick*" or finger-prick or "heel prick*" or heel-prick* or "heel lance*" or heel-lance* or "heel puncture*" or heel-puncture* or suture* or (laceration* n3 repair*))
S2. needles*
S1. (MH "Needles")

STUDY DESIGN FILTER FOR CINAHL

((MH "Random Assignment") or (MH "Single-Blind Studies") or (MH "Double-Blind Studies") or (MH "Triple-Blind Studies") or (MH "Crossover Design") or (MH "Factorial Design") or (MH "Placebos") or (MH "Clinical Trials")) or (multi centre study or multi center study or multi-centre study or multi-center study or cross-over or crossover or placebo* or random* or trial* or control* or group* or latin square)

WHAT'S NEW

Last assessed as up-to-date: 25 July 2011.

Date	Event	Description
12 April 2012	Amended	The tables originally placed in the Summary of Findings section of the review were moved to 'Additional tables' as they were identified not to be Summary of Findings tables. The Editorial office actioned this on behalf of the author

HISTORY

Protocol first published: Issue 4, 2006

Review first published: Issue 10, 2011

Date	Event	Description
10 November 2008	Amended	Further RevMan 5 conversion changes.
28 July 2008	Amended	Converted to new review format.

CONTRIBUTIONS OF AUTHORS

One review author (RPR) created the search strategy under the guidance of three Cochrane-affiliated Librarians (Elizabeth Ulryk, Sylvia Bickley, Caroline Struthers). One review author contributed to the modification of the strategies and protocol (RPR). Nine review authors (RPR, NR, KT, LU, RH, LDO, SAK, JHS, AG) independently screened titles and abstracts of studies from literature searches for inclusion in the review. Two review authors located and obtained articles (KT and AG). Six review authors performed data extraction (RPR, NR, RH, LDO, SAH, JHS). All review authors were involved reviewing the manuscript (RPR, NR, KT, LU, RH, LDO, SAK, JHS, BS, AG). Dr. Rebecca Pillai Riddell will be responsible for the update of this review.

DECLARATIONS OF INTEREST

None known

SOURCES OF SUPPORT

Internal sources

- Lillian Wright Foundation for Maternal-Child Health, York University, Canada.

Lillian Wright Foundation for Maternal-Child Health has also provided funds to: Ms. Nicole Racine, Ms. Laila Din Osmun, and Ms. Rachel Horton.

- York University, Canada.

Operating funds for this project were provided by start up funds from York University to Dr. Pillai Riddell.

External sources

- Canadian Institutes for Health Research, Canada.

Provided salary support to Dr. Pillai Riddell (Post-doctoral Fellowship, New Investigator Award) and doctoral fellowships to Ms. Laila Din Osmun and Ms. Jessica Hillgrove Stuart.

- Ontario Graduate Scholarship, Canada.

Provided support to Ms. Nicole Racine, Ms. Rachel Horton, and Ms. Sara Ahola Kohut.

- Social Sciences and Humanities Research Council, Canada.

Provided support to Ms. Nicole Racine.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

- Only behavioral measures were analyzed. Due to significant heterogeneity and missing data, physiological measures were not analyzed for our review.
- Despite our attempt to find studies examining postoperative pain, none of the studies included in our review were in a postsurgical context.
- Due to the complexity involved with a treatment X age X pain response analysis, the decision was made not to analyze studies by the fourth dimension (different types of outcome measures).
- No studies were located that could address the role of non-pharmacological studies as adjuvants to pharmacological studies.
- Age groups were collapsed into three groups, instead of four, based on the ages found in the located studies.
- The standardized mean difference was analyzed instead of the mean difference.
- The Yates Risk of Bias Scale was modified further to have the final question (Control Group) be out of 1 not 2.

INDEX TERMS

Medical Subject Headings (MeSH)

*Pain Management; Acute Disease; Infant Care [*methods]; Infant, Newborn; Infant, Premature; Needles [*adverse effects]; Pain [physiopathology]; Punctures [*adverse effects]; Randomized Controlled Trials as Topic; Sucking Behavior

MeSH check words

Child, Preschool; Humans; Infant