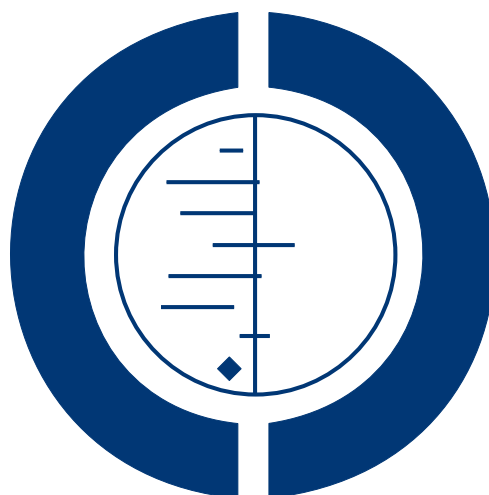


# Massage for promoting mental and physical health in typically developing infants under the age of six months (Review)

Bennett C, Underdown A, Barlow J



**THE COCHRANE  
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2013, Issue 4

<http://www.thecochranelibrary.com>



---

Massage for promoting mental and physical health in typically developing infants under the age of six months (Review)  
Copyright © 2013 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## TABLE OF CONTENTS

HEADER . . . . .	1
ABSTRACT . . . . .	1
PLAIN LANGUAGE SUMMARY . . . . .	2
BACKGROUND . . . . .	3
OBJECTIVES . . . . .	4
METHODS . . . . .	4
RESULTS . . . . .	7
Figure 1. . . . .	12
Figure 2. . . . .	13
DISCUSSION . . . . .	25
AUTHORS' CONCLUSIONS . . . . .	27
ACKNOWLEDGEMENTS . . . . .	28
REFERENCES . . . . .	28
CHARACTERISTICS OF STUDIES . . . . .	33
DATA AND ANALYSES . . . . .	74
Analysis 1.1. Comparison 1 Infant massage versus control - physical development, Outcome 1 Weight. . . . .	82
Analysis 1.2. Comparison 1 Infant massage versus control - physical development, Outcome 2 Weight: subgroup analyses (duration of intervention). . . . .	85
Analysis 1.3. Comparison 1 Infant massage versus control - physical development, Outcome 3 Length. . . . .	86
Analysis 1.4. Comparison 1 Infant massage versus control - physical development, Outcome 4 Length: subgroup analyses (duration of intervention). . . . .	88
Analysis 1.5. Comparison 1 Infant massage versus control - physical development, Outcome 5 Head circumference. . . . .	89
Analysis 1.6. Comparison 1 Infant massage versus control - physical development, Outcome 6 Head circumference: subgroup analyses (duration of intervention). . . . .	90
Analysis 1.7. Comparison 1 Infant massage versus control - physical development, Outcome 7 Mid arm circumference. . . . .	91
Analysis 1.8. Comparison 1 Infant massage versus control - physical development, Outcome 8 Mid leg/thigh circumference. . . . .	92
Analysis 1.9. Comparison 1 Infant massage versus control - physical development, Outcome 9 Abdominal circumference. . . . .	92
Analysis 1.10. Comparison 1 Infant massage versus control - physical development, Outcome 10 Chest circumference. . . . .	93
Analysis 1.11. Comparison 1 Infant massage versus control - physical development, Outcome 11 Hormones: cortisol. . . . .	94
Analysis 1.12. Comparison 1 Infant massage versus control - physical development, Outcome 12 Hormones: norepinephrine. . . . .	95
Analysis 1.13. Comparison 1 Infant massage versus control - physical development, Outcome 13 Hormones: epinephrine. . . . .	95
Analysis 1.14. Comparison 1 Infant massage versus control - physical development, Outcome 14 Hormones: serotonin. . . . .	96
Analysis 1.15. Comparison 1 Infant massage versus control - physical development, Outcome 15 Hormones: 6-sulphatoxymelatonin secretion. . . . .	97
Analysis 1.16. Comparison 1 Infant massage versus control - physical development, Outcome 16 Biochemical markers: Bilirubin (7 days PN). . . . .	97
Analysis 1.17. Comparison 1 Infant massage versus control - physical development, Outcome 17 Crying or fussing time. . . . .	98
Analysis 1.18. Comparison 1 Infant massage versus control - physical development, Outcome 18 Crying frequency (times). . . . .	99
Analysis 1.19. Comparison 1 Infant massage versus control - physical development, Outcome 19 Sleep/wake behaviours (Thoman). . . . .	100
Analysis 1.20. Comparison 1 Infant massage versus control - physical development, Outcome 20 Behavioural state immediately post-intervention (Thoman). . . . .	101
Analysis 1.21. Comparison 1 Infant massage versus control - physical development, Outcome 21 Sleep duration over 24hr period. . . . .	102
Analysis 1.22. Comparison 1 Infant massage versus control - physical development, Outcome 22 Mean increase in 24h sleep. . . . .	103
Analysis 1.23. Comparison 1 Infant massage versus control - physical development, Outcome 23 Mean increase in duration of night sleep. . . . .	104

Analysis 1.24. Comparison 1 Infant massage versus control - physical development, Outcome 24 Mean increase in duration of day sleep. . . . .	105
Analysis 1.25. Comparison 1 Infant massage versus control - physical development, Outcome 25 Mean increase in duration of first morning sleep after massage. . . . .	105
Analysis 1.26. Comparison 1 Infant massage versus control - physical development, Outcome 26 Sleep (total hours per night). . . . .	106
Analysis 1.27. Comparison 1 Infant massage versus control - physical development, Outcome 27 Number of naps (total number of naps). . . . .	107
Analysis 1.28. Comparison 1 Infant massage versus control - physical development, Outcome 28 Number of naps in day. . . . .	107
Analysis 1.29. Comparison 1 Infant massage versus control - physical development, Outcome 29 Number of naps at night. . . . .	108
Analysis 1.30. Comparison 1 Infant massage versus control - physical development, Outcome 30 Night Wake Frequency (times). . . . .	109
Analysis 1.31. Comparison 1 Infant massage versus control - physical development, Outcome 31 Night wake duration. . . . .	110
Analysis 1.32. Comparison 1 Infant massage versus control - physical development, Outcome 32 Blood flow (post intervention). . . . .	111
Analysis 1.33. Comparison 1 Infant massage versus control - physical development, Outcome 33 Formula intake. . . . .	112
Analysis 1.34. Comparison 1 Infant massage versus control - physical development, Outcome 34 Illness. . . . .	113
Analysis 1.35. Comparison 1 Infant massage versus control - physical development, Outcome 35 Illness and clinic visits. . . . .	114
Analysis 2.1. Comparison 2 Infant massage versus control - mental health and development, Outcome 1 Infant temperament meta-analyses. . . . .	115
Analysis 2.2. Comparison 2 Infant massage versus control - mental health and development, Outcome 2 Infant temperament (CCTI) post intervention. . . . .	116
Analysis 2.3. Comparison 2 Infant massage versus control - mental health and development, Outcome 3 Infant temperament (Infant behaviour questionnaire (IBQ) post intervention). . . . .	117
Analysis 2.4. Comparison 2 Infant massage versus control - mental health and development, Outcome 4 Infant temperament questionnaire (revised RITQ (Carey)) post-intervention 4 months. . . . .	118
Analysis 2.5. Comparison 2 Infant massage versus control - mental health and development, Outcome 5 Infant temperament questionnaire (revised RITQ (Carey)) follow-up 8 months. . . . .	120
Analysis 2.6. Comparison 2 Infant massage versus control - mental health and development, Outcome 6 Infant Care Questionnaire post-intervention. . . . .	122
Analysis 2.7. Comparison 2 Infant massage versus control - mental health and development, Outcome 7 Infant Care Questionnaire follow-up 1 year. . . . .	123
Analysis 2.8. Comparison 2 Infant massage versus control - mental health and development, Outcome 8 Infant attachment (Q set). . . . .	124
Analysis 2.9. Comparison 2 Infant massage versus control - mental health and development, Outcome 9 Child behaviour (HOME). . . . .	124
Analysis 2.10. Comparison 2 Infant massage versus control - mental health and development, Outcome 10 Eyberg Child Behaviour Inventory (ECBI) - Intensity domain. . . . .	125
Analysis 2.11. Comparison 2 Infant massage versus control - mental health and development, Outcome 11 Eyberg Child Behaviour Inventory (ECBI) - Problem domain. . . . .	126
Analysis 2.12. Comparison 2 Infant massage versus control - mental health and development, Outcome 12 Mother and child interaction meta-analysis - Total NCATS and Murray Global. . . . .	127
Analysis 2.13. Comparison 2 Infant massage versus control - mental health and development, Outcome 13 Nursing Child Feeding Assessment Scale (NCAFS) - Total. . . . .	128
Analysis 2.14. Comparison 2 Infant massage versus control - mental health and development, Outcome 14 Nursing Child Assessment Teaching Scale (NCATS) - Mother. . . . .	128
Analysis 2.15. Comparison 2 Infant massage versus control - mental health and development, Outcome 15 Nursing Child Assessment Teaching Scale (NCATS) - Child. . . . .	129
Analysis 2.16. Comparison 2 Infant massage versus control - mental health and development, Outcome 16 Maternal sensitivity - warm to cold (Murray). . . . .	130

Analysis 2.17. Comparison 2 Infant massage versus control - mental health and development, Outcome 17 Maternal sensitivity - non-intrusive to intrusive (Murray). . . . .	131
Analysis 2.18. Comparison 2 Infant massage versus control - mental health and development, Outcome 18 Maternal sensitivity - remoteness (Murray). . . . .	132
Analysis 2.19. Comparison 2 Infant massage versus control - mental health and development, Outcome 19 Infant interactions - infant performance - attentive to non attentive (Murray). . . . .	133
Analysis 2.20. Comparison 2 Infant massage versus control - mental health and development, Outcome 20 Infant interactions - lively to inert (Murray). . . . .	134
Analysis 2.21. Comparison 2 Infant massage versus control - mental health and development, Outcome 21 Infant interactions - happy to distressed (Murray). . . . .	135
Analysis 2.22. Comparison 2 Infant massage versus control - mental health and development, Outcome 22 Parenting stress (PSI Abidin) child characteristics subscale. . . . .	136
Analysis 2.23. Comparison 2 Infant massage versus control - mental health and development, Outcome 23 Psychomotor Development Indices (PDI) meta-analysis post-intervention. . . . .	137
Analysis 2.24. Comparison 2 Infant massage versus control - mental health and development, Outcome 24 Bayley Psychomotor Development Index (PDI) follow-up. . . . .	138
Analysis 2.25. Comparison 2 Infant massage versus control - mental health and development, Outcome 25 Mental Development Indices (MDI) meta-analysis post-intervention. . . . .	139
Analysis 2.26. Comparison 2 Infant massage versus control - mental health and development, Outcome 26 Bayley Mental Development Index (MDI) follow-up. . . . .	140
Analysis 2.27. Comparison 2 Infant massage versus control - mental health and development, Outcome 27 Gessel/Capital meta-analysis (post intervention). . . . .	141
Analysis 2.28. Comparison 2 Infant massage versus control - mental health and development, Outcome 28 Gessel Developmental Quotient (post intervention). . . . .	142
Analysis 2.29. Comparison 2 Infant massage versus control - mental health and development, Outcome 29 Capital institute Mental Checklist (post intervention). . . . .	143
Analysis 2.30. Comparison 2 Infant massage versus control - mental health and development, Outcome 30 Gessel Developmental Quotient (follow-up 6 months). . . . .	144
Analysis 2.31. Comparison 2 Infant massage versus control - mental health and development, Outcome 31 Attachment patterns (strange situation procedure). . . . .	145
Analysis 2.32. Comparison 2 Infant massage versus control - mental health and development, Outcome 32 Distractibility (toy) follow-up 1 year. . . . .	146
Analysis 2.33. Comparison 2 Infant massage versus control - mental health and development, Outcome 33 Habituation. . . . .	147
Analysis 2.34. Comparison 2 Infant massage versus control - mental health and development, Outcome 34 Seconds to habituation. . . . .	147
Analysis 2.35. Comparison 2 Infant massage versus control - mental health and development, Outcome 35 Trials to habituation. . . . .	148
Analysis 2.36. Comparison 2 Infant massage versus control - mental health and development, Outcome 36 Post habituation. . . . .	148
Analysis 2.37. Comparison 2 Infant massage versus control - mental health and development, Outcome 37 Habituation test. . . . .	149
ADDITIONAL TABLES . . . . .	149
APPENDICES . . . . .	151
WHAT'S NEW . . . . .	156
HISTORY . . . . .	156
CONTRIBUTIONS OF AUTHORS . . . . .	157
DECLARATIONS OF INTEREST . . . . .	157
SOURCES OF SUPPORT . . . . .	157
DIFFERENCES BETWEEN PROTOCOL AND REVIEW . . . . .	158
INDEX TERMS . . . . .	158

[Intervention Review]

# Massage for promoting mental and physical health in typically developing infants under the age of six months

Cathy Bennett<sup>1</sup>, Angela Underdown<sup>2</sup>, Jane Barlow<sup>2</sup>

<sup>1</sup>Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK. <sup>2</sup>Division of Mental Health and Wellbeing, Warwick Medical School, Coventry, UK

Contact address: Jane Barlow, Division of Mental Health and Wellbeing, Warwick Medical School, University of Warwick, Gibbett Hill Road, Coventry, CV4 7AL, UK. [jane.barlow@warwick.ac.uk](mailto:jane.barlow@warwick.ac.uk).

**Editorial group:** Cochrane Developmental, Psychosocial and Learning Problems Group.

**Publication status and date:** New search for studies and content updated (no change to conclusions), published in Issue 4, 2013.

**Review content assessed as up-to-date:** 20 December 2011.

**Citation:** Bennett C, Underdown A, Barlow J. Massage for promoting mental and physical health in typically developing infants under the age of six months. *Cochrane Database of Systematic Reviews* 2013, Issue 4. Art. No.: CD005038. DOI: 10.1002/14651858.CD005038.pub3.

Copyright © 2013 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## ABSTRACT

### Background

Infant massage is increasingly being used in the community with babies and their primary caregivers. Anecdotal reports suggest benefits for sleep, respiration and elimination, the reduction of colic and wind, and improved growth. Infant massage is also thought to reduce infant stress and promote positive parent-infant interaction.

### Objectives

The aim of this review was to assess whether infant massage is effective in promoting infant physical and mental health in low-risk, population samples.

### Search methods

Relevant studies were identified by searching the following electronic databases up to June 2011: CENTRAL; MEDLINE; EMBASE; CINAHL; PsycINFO; Maternity and Infant Care; LILACS; WorldCat (dissertations); ClinicalTrials.gov; China Masters' Theses; China Academic Journals; China Doctoral Dissertations; China Proceedings of Conference. We also searched the reference lists of relevant studies and reviews.

### Selection criteria

We included studies that randomised healthy parent-infant dyads (where the infant was under the age of six months) to an infant massage group or a 'no-treatment' control group. Studies had to have used a standardised outcome measure of infant mental or physical development.

### Data collection and analysis

Mean differences (MD) and standardised mean differences (SMD) and 95% confidence intervals (CIs) are presented. Where appropriate, the results have been combined in a meta-analysis using a random-effects model.

## Main results

We included 34 studies, which includes one that was a follow-up study and 20 that were rated as being at high risk of bias.

We conducted 14 meta-analyses assessing physical outcomes post-intervention. Nine meta-analyses showed significant findings favouring the intervention group for weight (MD -965.25 g; 95% CI -1360.52 to -569.98), length (MD -1.30 cm; 95% CI -1.60 to -1.00), head circumference (MD -0.81 cm; 95% CI -1.18 to -0.45), arm circumference (MD -0.47 cm; 95% CI -0.80 to -0.13), leg circumference (MD -0.31 cm; 95% CI -0.49 to -0.13), 24-hour sleep duration (MD -0.91 hr; 95% CI -1.51 to -0.30), time spent crying/fussing (MD -0.36; 95% CI -0.52 to -0.19), decreased levels of blood bilirubin (MD -38.11 mmol/L; 95% CI -50.61 to -25.61), and there were fewer cases of diarrhoea, RR 0.39; 95% CI 0.20 to 0.76). Non-significant results were obtained for cortisol levels, mean increase in duration of night sleep, mean increase in 24-hour sleep and for number of cases of upper respiratory tract disease and anaemia.

Sensitivity analyses were conducted for weight, length and head circumference, and only the finding for length remained significant following removal of studies judged to be at high risk of bias. These three outcomes were the only ones that could also be meta-analysed at follow-up; although both weight and head circumference continued to be significant at 6-month follow-up, these findings were obtained from studies conducted in Eastern countries only. No sensitivity analyses were possible.

We conducted 18 meta-analyses measuring aspects of mental health and development. A significant effect favouring the intervention group was found for gross motor skills (SMD -0.44; 95% CI -0.70 to -0.18), fine motor skills (SMD -0.61; 95% CI -0.87 to -0.35), personal and social behaviour (SMD -0.90; 95% CI -1.61 to -0.18) and psychomotor development (SMD -0.35; 95% CI -0.54 to -0.15); although the first three findings were obtained from only two studies, one of which was rated as being at high risk of bias, and the finding for psychomotor development was not maintained following removal of studies judged to be at high risk of bias in a sensitivity analysis. No significant differences were found for a range of aspects of infant temperament, parent-infant interaction and mental development. Only parent-infant interaction could be meta-analysed at follow-up, and the result was again not significant.

## Authors' conclusions

These findings do not currently support the use of infant massage with low-risk groups of parents and infants. Available evidence is of poor quality, and many studies do not address the biological plausibility of the outcomes being measured, or the mechanisms by which change might be achieved. Future research should focus on the impact of infant massage in higher-risk groups (for example, demographically and socially deprived parent-infant dyads), where there may be more potential for change.

## PLAIN LANGUAGE SUMMARY

### Message for promoting mental and physical health in infants under the age of six months

This review aimed to assess the impact of infant massage on mental and physical outcomes for healthy mother-infant dyads in the first six months of life. A total of 34 randomised trials were included. Twenty of these had significant problems with their design and the way they were carried out. This means that we are not as confident as we would otherwise be that the findings are valid. That is to say, the findings of these 20 included studies may over- or under-estimate the true effect of massage therapy.

We combined the data for 14 outcomes measured physical health and 18 outcomes measured aspects of mental health or development. The results show limited statistically significant benefits for a number of aspects of physical health (for example, weight, length, head/arm/leg circumference, 24-hour sleep duration; time spent crying or fussing; blood bilirubin and number of episodes of illness) and mental health/development (for example, fine/gross motor skills personal and social behaviour and psychomotor development). However, all significant results were lost either at later follow-up points or when we removed the large number of studies regarded to be at high risk of bias.

These findings do not currently support the use of infant massage with low-risk population groups of parents and infants. The results obtained from this review may be due to the poor quality of many of the included studies, the failure to address the mechanisms by which infant massage could have an impact on the outcomes being assessed, and the inclusion of inappropriate outcomes for population groups (such as weight gain). Future research should focus on the benefits of infant massage for higher-risk population groups (for example, socially deprived parent-infant dyads), the duration of massage programmes, and could address differences between babies being massaged by parents or healthcare professionals.

## BACKGROUND

In many areas of the world, especially in the African and Asian continents, indigenous South Pacific cultures and the Soviet Union, infant massage is a traditional practice (Field 1996b). A survey of 332 primary caregivers of neonates in Bangladesh, for example, found that 96% engaged in massage of the infant's whole body between one and three times daily (Darmstadt 2002a).

In Western cultures, infant massage was initially used to improve outcomes for infants in neonatal intensive care units (NICUs) where the environment can be stressful for infants, and where tactile stimulation can be poor (Vickers 2004). Developing understanding about the importance for infant development of warm, sensitive, attentive interactions (see Tronick 2007 for an overview), 'midrange' responsiveness on the part of the primary caregiver (that is, compared with heightened or lowered responsiveness) (Beebe 2010) and body-based interactions (Shai 2011) (see below [Description of the intervention](#) for further detail), has resulted in an increased interest in the possible role of infant massage to support early sensitive parent-infant relationships, particularly where the mother may be experiencing difficulties such as postnatal depression (Kersten-Alvarez 2011).

The practice of infant massage varies across the world with western cultures adapting some of the traditional practices from Eastern cultures. However, there is considerable variability in the techniques being promoted, with the International Association of Infant Massage teaching the use of nurturing touch and respectful communication, while other schools of training emphasise yoga-based movements and flexibility (Underdown 2011).

### Description of the intervention

#### Physiological and psychological impact of infant massage

Reviews of the effectiveness of infant massage have to date focused on preterm infants, and outcomes that are important in this group, including weight gain, activity levels and length of stay in hospital (Ireland 2000; Vickers 2004). Although Vickers 2004 found that massage improved daily weight gain in preterm infants by 5.1 g (95% CI 3.5 to 6.7), including some evidence of a small positive effect on weight at four to six months, and reduced the length of hospital stay by 4.5 days (95% CI 2.4 to 6.5), concerns were raised about the methodological quality of the included studies, particularly in respect of selective reporting of outcomes. Ireland 2000 also showed a beneficial effect of infant massage on weight gain, activity level and hospital stay. Studies that have examined the impact of infant massage on other high-risk groups, such as women experiencing postnatal depression, have found evidence of impact on maternal sensitivity (Kersten-Alvarez 2011).

The potential role of interventions such as infant massage even with groups of parents not at high risk has been highlighted by recent research in the field of developmental psychology and infant mental health, which has indicated the importance of parental attuned and sensitive caregiving for infant attachment security. Parental sensitivity to an infant's signals and cues at two months has been shown to be associated with secure attachment status at nine months (De Wolff 1997); and low sensitivity shown to be associated with compromised cognitive and emotional development (Murray 1992), and behavioural and physiological difficulties (Gianino 1988; Tronick 2007; Degnan 2008). The quality of the parent-infant interaction relies to a large extent on the parent's ability to read and respond appropriately to the infant's emotional state (Kropp 1987; Zeanah 2000).

The potential importance of 'dyadic' and body-based approaches such as infant massage have also been emphasised by developments in the field of infant mental health that have focused attention on the importance of dyadic states of consciousness (Tronick 2007), and parent-infant communication as a bi-directional, moment-to-moment process occurring across multiple modalities (Beebe 2010), in addition to the importance of whole-body kinaesthetic patterns during parent-infant interactions (Shai 2011).

Tronick 1989 developed the Mutual Regulation Model to refer to the 'dyadic system of regulation and communication in which the caregiver and infant mutually regulate the physiological and emotional states of the other'. This model postulates that infants have a range of self-organising neuro-behavioural capacities that are used to organise both behavioural states and a range of biopsychological processes (for example, self-regulation of arousal, selective attentional learning and memory, social engagement etc) (Tronick 2007, p.8). It also postulates that the sensitive caregiver helps the infant to regulate these states by being attuned to the infant's 'organised communicative displays' that indicate their internal state (Tronick 2007, p.10). Tronick's research identified the bi-directional, synchronous and co-ordinated nature of mother-infant interaction, in which sensitive caregivers are able to repair mismatched states (Tronick 1982). His research using the 'Still-Face Perturbation' with mothers experiencing postnatal depression, identified the significant impact of disturbances to the communicative regulatory system in terms of its role in the intergenerational transfer of mood (see, for example, Tronick 2007 for a summary).

The Dyadic Systems Approach of Beebe 2010 has broadened the focus of parental regulation of infant emotional distress to include recognition of the importance of multiple communication modalities including affect (facial and vocal); visual attention (gaze on/off), touch (maternal touch, infant initiated touch); spatial orientation (mother orientation from sitting upright to leaning forward to looming in; infant head orientation from face-to-face to arch), alongside a composite variable of facial-visual engagement (Beebe 2010, p.9). This approach recognises that different modalities can convey discordant information that can be difficult for

the infant to co-ordinate, and that may be the basis of later problems such as ‘disorganised attachment’ (Beebe 2010). Beebe 2010 showed that dyadic interaction of future insecurely attached (that is, ‘resistant’) infants was characterised by dysregulated tactile and spatial exchanges, generating approach-withdrawal patterns, while the interaction of future ‘disorganised’ infants was characterised by intrapersonal and interpersonal discordance or conflict in the face of intense infant distress (Beebe 2010, p.6-7).

Similarly, recent attempts to operationalise the concept of ‘mentalisation’ (Fonagy 2002; Fonagy 2007), which emphasises the importance of the parent’s ability to reflect on their infant’s internal states for later secure attachment (Arnott 2007), have resulted in the development of the concept of Parental Embodied Mentalisation (PEM). PEM refers explicitly to the quality of dynamic moment-to-moment changes in whole-body kinaesthetic patterns during parent-infant interactions (Shai 2011), and focuses on the parents’ capacity to ‘a) implicitly conceive, comprehend, and extrapolate the infant’s mental states (such as wishes, desires or preferences) from the infant’s whole-body kinaesthetic expressions; and b) adjust one’s own kinaesthetic patterns accordingly’ (ibid, p.175). The focus of PEM is on ‘how’ interactive bodily actions are performed rather than ‘what’ actions are performed, and as such includes both spatial and temporal dynamic contours. As with the work of Beebe 2010, this approach treats the ‘dyad’ as the unit of action, and the moment-to-moment exchanges as being bi-directional in terms of their mutual influence. There is also recognition of the importance of interactive repair following rupture to interactive synchrony, but with a particular focus on the parent’s contribution in terms of their kinaesthetic adjustment.

The importance of identifying effective methods of supporting early parenting is also indicated by evidence about the prevalence of problems such as sleep, colic, excessive crying and stress (Keren 2001), which have been shown to be associated with the parent-infant relationship (Papousek 1995), alongside their impact on the child’s later development including delays in motor, language and cognitive development at three years of age (Degangi 2000).

### How the intervention might work

Some of the mechanisms by which massage might promote improved outcomes in infants have been investigated in both animal and human populations. For example, in rodents high frequency of licking and grooming of the pups has been shown to be associated with reduced fearfulness and dampened responsiveness to stress in adulthood as a result of such stimulation on the hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal reactivity (Liu 1997). Other studies have shown that higher frequency licking and grooming is associated with improved cognitive development in rats (specifically greater spatial learning and memory performance) (Liu 2002), as a result of enhanced synaptogenesis and neuronal survival in the hippocampus (Bredy 2003).

A number of studies have examined the potential mechanisms by which tactile stimulation could impact on human infants. For example, Field 1996b found that infant massage resulted in reduced catecholamine (norepinephrine and epinephrine) and cortisol excretion, and it is now recognised that high cortisol levels have damaging effects on the developing brain, particularly in terms of the later capacity of such infants to regulate their stress levels (Gunnar 1998; Gunnar 2007). Another study reported an effect on release of melatonin (6-sulphatoxymelatonin), which is involved in the adjustment of circadian rhythms and sleep (Ferber 2002), and Uvnas-Moberg 1987 reported that massage increased vagal activity and secretion of insulin and gastrin improving the absorption of food, and thereby suggesting a plausible biological mechanism for the impact of infant massage on growth (Vickers 2004).

### Why it is important to do this review

Increasing evidence about the importance of early relationships for optimal infant development has resulted in a drive to find acceptable effective interventions to support early interaction in both high-risk and population groups. The effectiveness of infant massage has been reviewed for a number of high-risk populations (for example, preterm infants; postnatally depressed women), and there is now a need to examine its effectiveness for population groups (that is, where there is has been no risk identified), in terms of both physical and mental health outcomes.

## OBJECTIVES

To assess whether infant massage is effective in promoting infant mental health, parent-infant interaction, or physical aspects of development in population samples of babies

## METHODS

### Criteria for considering studies for this review

#### Types of studies

Studies were included if participants had been randomised to either an infant massage group or a control group that received no intervention. The review also included quasi-randomised study designs.



## Types of participants

Babies under the age of six months were eligible for inclusion. Studies focusing on preterm and low birthweight babies receiving massage within a hospital setting were excluded.

## Types of interventions

Studies were included if they evaluated the effectiveness of infant massage, irrespective of the theoretical basis or cultural practice underpinning the massage. Infant massage was defined in this review as systematic tactile stimulation by human hands. This included studies where the technique of infant massage had been specifically taught to parents and/or staff, and evaluations of infant massage where it was used as a routine cultural practice. Multimodal interventions, of which massage was a part, were only included if the benefits of massage as a separate intervention could be elicited.

## Types of outcome measures

To be eligible for inclusion in the review, studies had to include at least one standardised instrument measuring the effect of infant massage on either infant mental health (for example, the CARE-Index to measure infant-adult interaction) or on physical health (for example, growth monitoring).

## Primary outcomes

### Physical outcomes

Weight and length; head, leg, arm, chest, abdominal circumference; illness and clinic visits/service use; hormone (for example, cortisol, epinephrine, norepinephrine, melatonin, serotonin) levels and blood flow; behavioural states (for example, sleep, wake and crying durations); formula intake.

### Mental and development outcomes

Infant temperament (for example, activity, soothability, emotionality and sociability etc); attachment; behaviour (for example, Eyberg Child Behaviour Inventory (ECBI); Nursing Child Teaching Assessment Scales (NCATS)); parent-infant interaction; development (for example, Bayley Scales); IQ (for example, Capital Institute Mental Checklist (China)).

### Timing of outcome measures

Post-intervention: immediately following the completion of the intervention.

Follow-up: between six and 12 months after the completion of the intervention.

## Search methods for identification of studies

### Electronic searches

The original search strategies are presented in [Appendix 1](#). For the updated review, the following databases were searched from 2005 onwards with the exception of *Maternity and Infant Care*, which was new for the updated review and therefore searched for all years. Searches for the updated review were run in May 2010 and updated in June 2011. The same search terms were used in both sets of searches (see [Appendix 3](#); [Appendix 4](#)).

- Cochrane Central Register of Controlled Trials (CENTRAL), 2011, Issue 3, last searched 20 June 2011
- Ovid MEDLINE, 1948 to June Week 2 2011, last searched 20 June 2011
- EMBASE, 1980 to 2011 Week 24, last searched 20 June 2011
- CINAHL, 1937 to current, last searched 20 June 2011
- PsycINFO, 1887 to current, last searched 20 June 2011
- *Maternity and Infant Care*, 1971 to June 2011, last searched 20 June 2011
- LILACS, last searched 20 June 2011
- WorldCat (limited to theses), last searched 20 June 2011
- ClinicalTrials.gov, searched 20 June 2011

The following four databases were searched for the update via the China Knowledge Resource Integrated Database (CNKI)

- China Masters' Theses, 2000 to current, searched 15 June 2011
- China Academic Journals, 1915 to current, searched 15 June 2011
- China Doctoral Dissertations, 1999 to current, searched 15 June 2011
- China Proceedings of Conference, searched 15 June 2011

We designed searches with the support of the Cochrane CD-PLPG group. The search terms were adapted for use in different databases. No methodological terms were included to ensure that all relevant papers were retrieved. There was no language restriction. Relevant papers were translated or data extracted by researchers fluent in written Chinese where necessary. For the update in 2011, we used a machine translation service (Google translate) to obtain details from studies written in languages other than English. Because automatically machine-generated translations are not necessarily accurate enough for the scientific purpose, we confirmed details with study investigators where possible.

### Searching other resources

Reference lists of articles identified through database searches and bibliographies of systematic and non-systematic review articles were examined to identify further relevant studies.

## Data collection and analysis

### Selection of studies

Titles and abstracts of trials identified through searches of electronic databases were independently screened by two review authors to determine whether they met the inclusion criteria (AU and JB; and VC and JH for Chinese studies). Abstracts that did not meet the inclusion criteria were rejected. Two independent review authors (AU and JB; and VC and YH for Chinese studies) assessed full copies of papers that appeared to meet the inclusion criteria. Uncertainties concerning the appropriateness of studies for inclusion in the review were resolved through consultation with a third review author (SSB). For the update in 2011, CB identified additional studies obtained by electronic searches and these were referred to JB and AU for a decision about whether they met the inclusion criteria of the review.

### Data extraction and management

Two review authors (AU and CB) independently extracted data and any queries were referred to JB. Data were entered into Review Manager 5 software (RevMan 5.1.7). Where data were not available in the published trial reports, we contacted study investigators to supply missing information.

### Assessment of risk of bias in included studies

In the previous published version of this review (Underdown 2006), two review authors (AU and JB) carried out the critical appraisal of the included studies. Disagreement was resolved by consultation with a third review author (SSB). Consistent with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), this version of the review incorporates additional elements into 'Risk of bias' tables that were not present in the previous published review. 'Risk of bias' assessments for the new included studies were carried out by CB and AU or JB. Differences were resolved by consensus. CB, JB and AU reassessed the study quality for the old included studies using the 'Risk of bias' assessment tool (Higgins 2011).

Risk of bias was assessed for each trial using the following criteria: sequence generation, allocation concealment, blinding of participants, personnel and outcome assessors, incomplete outcome data and whether there was any assessment of the distribution of confounders. Where there was insufficient information in the trial report to make a judgement, and the study was published less than 10 years previously, we contacted trial investigators for further information.

### Measures of treatment effect

Continuous outcomes were analysed if the mean and standard deviation of endpoint measures were presented. Where mean scores

were not available, we presented significance levels reported in the paper. Where baseline or pre-treatment means were available, these were examined to determine similarities between groups. For the meta-analyses of continuous outcomes, we estimated mean differences (MDs) between groups. In the case of continuous outcome measures where data were reported on different and incompatible scales, we analysed data using the standardised mean difference (SMD). We presented the SMD and 95% confidence intervals (CIs) for individual outcomes in individual studies. The SMD was calculated by dividing the MD in post-intervention scores between the intervention and control groups by the pooled standard deviation.

Where it was not possible to synthesise the data, we present effect sizes and 95% CIs for individual outcomes in each study.

One study compared four different types of massage oil with outcomes for a control group (Argawal 2000). In order to incorporate the results of this study, we calculated a pooled estimate of outcomes across the four treatment groups.

### Unit of analysis issues

Randomisation of clusters can result in an overestimate of the precision of the results (with a higher risk of a Type I error) where their use has not been compensated for in the analysis. None of the included studies employed cluster randomisation.

For studies where there was more than one active intervention and only one control group, we selected the intervention that most closely matched our inclusion criteria and excluded the others. (Chapter 16.5.4, Higgins 2011).

In (Argawal 2000), where all four intervention groups employed massage (with different oils), we combined the groups to create a single pair-wise comparison. In practice, we combined the data from the massage groups to produce a pooled mean and SD.

### Dealing with missing data

Where data were not available in the published trial reports or clarification was needed, we contacted trial investigators to supply missing information. It should be noted that one of the limitations of this approach is that it assumes independence of comparisons, and ignores the dependency from sharing the same control group.

### Assessment of heterogeneity

An assessment was made of the extent to which there were variations in the methods, populations, interventions or outcomes. Consistency of results was assessed by visual inspection of the forest plot and by examining  $I^2$  (Higgins 2002), a quantity which describes the approximate proportion of variation in point estimates that is due to heterogeneity rather than sampling error. We supplemented this with a test of homogeneity to determine the strength of evidence that the heterogeneity was genuine. The possible reasons for heterogeneity were explored by scrutinising the studies and, where appropriate, by performing subgroup analyses.

There was some clinical heterogeneity across the included studies (see [Description of studies](#)), and also some statistical heterogeneity for the small number of outcomes for which it was possible to combine the data. Quantitative syntheses of the data have therefore been undertaken using a random-effects model.

### Data synthesis

Where appropriate, we used meta-analyses to combine comparable outcome measures across studies, using a random-effects model.

### Subgroup analysis and investigation of heterogeneity

In the updated review we made a post hoc decision to investigate the effect of the duration of intervention on outcome. We categorised the duration of the massage programmes as follows: brief (a single session); short-term where the intervention took place for up to four weeks; medium-term where the intervention took place for at least four weeks and up to 12 weeks; and long-term where the intervention took place for more than 12 weeks. We did not carry out further subgroup analyses such as a comparison of massage provider. This decision is discussed further ([Discussion](#)).

### Sensitivity analysis

A sensitivity analysis was used to assess the robustness of the findings by examining the impact of one large study ([Kim 2003](#)). This was undertaken because we were concerned about the level of heterogeneity produced by this meta-analysis, and that the results of this study were influenced by the fact that, compared with the other included studies, the sample comprised infants receiving unusually low levels of tactile stimulation as a result of being in an orphanage.

In this updated review, we made a post-hoc decision, based on the clinical and statistical heterogeneity of the included studies, to perform sensitivity analyses based on the geographical location of the studies (East or West) and study quality (high risk of bias due to inadequate randomisation).

## RESULTS

### Description of studies

See: [Characteristics of included studies](#); [Characteristics of excluded studies](#).

### Results of the search

In 2005, for the original published version of the review, we reviewed 809 abstracts from international databases; most were of no relevance to full-term infants. After closer inspection of 35

abstracts, nine studies were identified as being suitable for inclusion ([Koniak-Griffin 1988](#); [Field 1996](#); [Cigales 1997](#); [Jump 1998](#); [Argawal 2000](#); [Onozawa 2001](#); [Elliott 2002](#); [Ferber 2002](#); [Kim 2003](#)); one other included study ([Koniak-Griffin 1995](#)) was a follow-up report of [Koniak-Griffin 1988](#). For this update, the follow-up report ([Koniak-Griffin 1995](#)) has been added to [Koniak-Griffin 1988](#). A handsearch of references was conducted, which resulted in the identification of one further study ([Ke 2001](#)). Of the 100+ abstracts reviewed from the Chinese databases, 12 studies were identified as suitable for inclusion ([Wang 1999](#); [Zhai 2001](#); [Duan 2002](#); [Shi 2002](#); [Sun 2004](#); [Xua 2004](#); [Ye 2004](#); [Liu CL 2005](#); [Liu DY 2005](#); [Lu 2005](#); [Na 2005](#); [Shao 2005](#)); one further study ([Liu 2001](#)) was assigned to the 'Awaiting assessment' category as further details could not be obtained at that time. This study ([Liu 2001](#)) was translated and included in the current update and is a report of two studies on infants of either birth to two months or three to six months of age. For the purposes of this review we treated this report as two individual studies ([Liu C 2001 0 to 2 months](#); [Liu C 2001 3 to 6 months](#)).

The update searches yielded 2179 hits in May 2010 and 1124 hits in June 2011. From closer inspection of 24 abstracts, we identified eight new studies that met the inclusion criteria: six studies from international databases ([Jing 2007](#); [Oswalt 2007](#); [Arikan 2008](#); [Narenji 2008](#); [O'Higgins 2008](#); [White-Traut 2009](#)) and two from Chinese databases ([Wang 2001](#); [Maimaiti 2007](#)). We searched the bibliography lists of all the new included studies and identified another two studies to include ([Cheng 2004](#); [Zhu 2010](#)).

In the previous version of the review published in 2006, 23 studies were included ([Underdown 2006](#)). In this updated version there are 34 included studies, of which 12 are new ([Liu C 2001 0 to 2 months](#); [Liu C 2001 3 to 6 months](#); [Wang 2001](#); [Cheng 2004](#); [Jing 2007](#); [Maimaiti 2007](#); [Oswalt 2007](#); [Arikan 2008](#); [Narenji 2008](#); [O'Higgins 2008](#); [White-Traut 2009](#); [Zhu 2010](#)).

### Included studies

#### Design

All 34 included studies were randomised parallel group trials. Four studies ([Argawal 2000](#); [Jing 2007](#); [Oswalt 2007](#); [Narenji 2008](#)) used a random number table to assign participants to intervention or control groups. [Elliott 2002](#) used a repeated measures design involving a randomised two-way layout with treatment factors 'carrying' and 'massage' as two levels to ensure that every dyad had an equal chance of being assigned to one of four groups. Nine studies were quasi-randomised ([Field 1996](#); [Jump 1998](#); [Zhai 2001](#); [Kim 2003](#); [Lu 2005](#); [Shao 2005](#); [O'Higgins 2008](#); [White-Traut 2009](#); [Zhu 2010](#)).

In five studies, insufficient details were provided to determine the exact method of randomisation ([Koniak-Griffin 1988](#); [Cigales 1997](#); [Onozawa 2001](#); [Ferber 2002](#); [Arikan 2008](#)).

In the remaining 15 studies, described in the study report as randomised (Wang 1999; Ke 2001; Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Wang 2001; Duan 2002; Shi 2002; Cheng 2004; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Liu DY 2005; Na 2005; Maimaiti 2007), insufficient details were provided to be certain that the study was in fact randomised and we were unable to obtain further details from the trial investigators.

In all 34 included studies, massage interventions were compared with normal care.

Five studies compared more than one intervention. Argawal 2000 compared four types of massage oil with a 'no treatment' control group: because the massage interventions were similar, we used pooled data from the three intervention groups. Arikan 2008 investigated massage, sucrose solution, herbal tea and infant formula versus control; we compared the massage and control groups. Elliott 2002 compared massage, supplemental carrying, both massage and supplemental carrying groups with a no treatment control group. We compared the massage group and the control group. Koniak-Griffin 1988 employed a four-arm design of massage only, massage combined with multisensory stimulation, or multisensory stimulation only, and a no treatment control group. We compared the unimodal massage intervention with the control group. White-Traut 2009 compared tactile only, auditory, tactile, visual, vestibular (ATVV) intervention with control. We compared the ATVV and control groups.

Jing 2007 used a massage and motion training intervention versus control. We included this study because motion training is integral to the Johnson massage method (Johnson 2011).

One study where the control group received rocking (Elliott 2002) was also included as this was considered to be usual soothing behaviour. The remaining studies compared massage with control (that is, no massage intervention or care as usual).

### Sample sizes

Thirty-four studies randomised 3984 participants. The largest study was Ke 2001 with 400 participants randomised; the smallest were Ferber 2002 (n = 21); Oswalt 2007 (n = 25) and White-Traut 2009 (n = 26).

### Participants

The infant participants were full-term babies of either sex, age six months or younger, with no underlying health conditions other than colic (Arikan 2008). The intervention commenced with newborn babies within one week of birth in Koniak-Griffin 1988; Wang 1999; Ke 2001; Wang 2001; Zhai 2001; Duan 2002; Elliott 2002; Ferber 2002; Shi 2002; Cheng 2004; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Liu DY 2005; Lu 2005; Na 2005; Shao 2005; Jing 2007; Maimaiti 2007; White-Traut 2009; Zhu 2010. Kim 2003 randomised participants within 14 days of birth.

Slightly older babies were studied in Argawal 2000 (six weeks of age); Arikan 2008 (2.29 months intervention; 2.28 months

in control); Cigales 1997 (four months old); Field 1996 (one to three months old infants); Jump 1998 (under nine months of age, mean age under six months); from birth to two months of age in Liu C 2001 0 to 2 months; and from three months of age to six months in Liu C 2001 3 to 6 months; Narenji 2008 (two months); O'Higgins 2008 and Onozawa 2001 (nine weeks of age); Oswalt 2007 (intervention 52.71 days; control 84 days). Koniak-Griffin 1988 reported follow-up results at 24 months post birth.

Mothers were diagnosed with depression in Field 1996 (adolescents); Onozawa 2001 (adults), or with depressive symptoms in O'Higgins 2008. In Oswalt 2007, the mothers were adolescents. One of the included studies focused on orphanage infants (Kim 2003). This study was included because there was no indication in the paper that the infants were not healthy full-term babies.

### Setting

Two studies (Cigales 1997; White-Traut 2009) were conducted in maternity hospital settings.

Nine studies (Koniak-Griffin 1988; Jump 1998; Argawal 2000; Onozawa 2001; Elliott 2002; Ferber 2002; Arikan 2008; Narenji 2008; O'Higgins 2008) were conducted in a community setting after training parents to carry out massage. A further seven studies (Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Wang 2001; Cheng 2004; Jing 2007; Maimaiti 2007; Zhu 2010) that were carried out in China were also undertaken by parents in the community after initial training in massage techniques.

Oswalt 2007 was set within a school-based parent training programme for adolescent mothers. One study (Field 1996) was conducted in a day-care centre. Kim 2003 was conducted in an orphanage.

In 13 studies (Wang 1999; Ke 2001; Zhai 2001; Duan 2002; Shi 2002; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Liu DY 2005; Lu 2005; Na 2005; Shao 2005), the setting was unclear, other than that the intervention took place in China; we were unable to obtain further information.

### Country

Studies were carried out in eight countries. In the West: UK (Onozawa 2001; O'Higgins 2008) and the USA (Koniak-Griffin 1988; Field 1996; Cigales 1997; Jump 1998; Elliott 2002; Oswalt 2007; White-Traut 2009) and Canada (Elliott 2002). In the East: Korea (Kim 2003), Israel (Ferber 2002), India (Argawal 2000), Iran (Narenji 2008) and Turkey (Arikan 2008). The remaining 20 studies were carried out in China (Wang 1999; Ke 2001; Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Wang 2001; Zhai 2001; Duan 2002; Shi 2002; Cheng 2004; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Liu DY 2005; Lu 2005; Na 2005; Shao 2005; Jing 2007; Maimaiti 2007; Zhu 2010).

## Interventions

### Massage provider

In four studies massage was offered by researchers (Field 1996; Cigales 1997; Kim 2003; White-Traut 2009). Kim 2003 involved orphans receiving a multimodal intervention of massage, talking and eye contact from research associates who were trained to be responsive to the infant's responses. White-Traut 2009 used trained researchers to deliver either a multimodal form of massage including auditory, tactile, visual and vestibular stimulation (ATVV) or tactile only stimulation (that is, we only included the ATVV group). Although it was not possible to isolate the effects of eye contact and talking, we included these studies because these components are an intrinsic part of some included infant massage programmes. Field 1996 used trained researchers to massage the infants of depressed adolescent mothers.

In seven studies massage was provided by the parent following instruction (Koniak-Griffin 1988; Argawal 2000; Elliott 2002; Ferber 2002; Jing 2007; Arikan 2008; Narenji 2008), and involved parents being taught massage techniques prior to them conducting massage on their infants in the home. Arikan 2008 trained mothers in massage providing them with an illustrated brochure with techniques. Argawal 2000 provided participating mothers with instruction and training, and their technique was monitored each week when they attended clinic to collect more oil. Elliott 2002 taught mothers the massage strokes when their infants were between seven and 10 days old, and a research assistant visited the home to monitor the parents' use of the technique. Parents also received an instructional videotape and written guidance. Ferber 2002 instructed mothers how to massage their infants as part of the bedtime routine and a research assistant telephoned on three occasions to ensure compliance. Jing 2007 trained parents using instruction, manuals and videos. Koniak-Griffin 1988 instructed mothers how to massage their infants and the massage technique was monitored using maternal self-report. Narenji 2008 instructed mothers to massage their babies with sesame oil, using a specific set of movements. O'Higgins 2008 invited mothers to attend a weekly massage class run by trained members of the International Association of Infant Massage (IAIM). Each group began with a discussion then focused on massage strokes as demonstrated by the instructors and on paying attention to infant cues. Oswalt 2007 trained the mothers in a class, each training session lasting approximately 30 minutes, the mothers also received a booklet illustrated with diagrams of the massage strokes and were asked to massage their infants daily for two months. Onozawa 2001 taught massage, and appropriate response to infant cues during massage, using trained IAIM instructors.

In seven of the 20 studies carried out in China, the massage was mostly administered by a nurse or member of the medical staff with specialist training in infant massage, following which the technique was taught to the parents who continued the massage

at home (Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Wang 2001; Cheng 2004; Jing 2007; Maimaiti 2007; Zhu 2010), although in Liu DY 2005 the intervention was apparently carried out throughout the 42-day intervention period by nurses. In the remaining 12 studies, also carried out in China, (Wang 1999; Ke 2001; Zhai 2001; Duan 2002; Shi 2002; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Lu 2005; Na 2005; Shao 2005), it was unclear from the published report who provided the massage intervention, and we were unable to obtain further details from the trial investigators.

### Dose and duration of intervention

The massage programmes evaluated in the included studies varied greatly in terms of duration and frequency. We categorised the duration of the intervention as brief (a single session), short-term (where the intervention took place for up to four weeks), medium-term (where the intervention took place for at least four weeks and up to 12 weeks) and long-term (where the intervention took place for at least 12 weeks and continued for up to 26 weeks).

We categorised two studies as brief interventions. In one study, infants were massaged once only for eight minutes (Cigales 1997); massage was administered only once prior to the conduct of an experimental task to assess the impact of massage on cognitions. In White-Traut 2009, infants received one 15-minute session of massage before collection of cortisol samples.

Ten studies were categorised as short-term interventions: in Arikan 2008, infants were given massage twice a day for 25 minutes during symptoms of colic for one week only. In another, infants received a daily 30-minute intervention over 14 days (Ferber 2002). In the Jump 1998 study, mothers and infants attended group sessions on a weekly basis for 45 to 60 minutes over the course of four weeks. During this time mothers were taught the massage techniques and were also given information about infant development. In the Kim 2003 study, infants were massaged for 15 minutes, twice daily for four weeks. In Narenji 2008, mothers massaged their infants twice daily for 10 minutes for four weeks (starting the massage just before morning and night sleep times). In Argawal 2000, infants received 10 minutes of massage daily over a four-week period. In Zhai 2001, Na 2005, Shao 2005 and Shi 2002 (all conducted in China), infants were massaged for 15-minute periods up to three times a day over a period extending up to 30 days.

Nineteen of the studies (Koniak-Griffin 1988; Field 1996; Wang 1999; Onozawa 2001; Ke 2001; Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Wang 2001; Duan 2002; Cheng 2004; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Lu 2005; Liu DY 2005; Oswalt 2007; O'Higgins 2008; Zhu 2010), delivered the intervention over an medium-term duration (from one month to up to three months). In the Field 1996 study, infants received 15 minutes of massage twice weekly over a period of six weeks and in the Koniak-Griffin 1988, study infants received five to seven minutes of massage once daily over three months. In two studies (Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months), massage

was delivered two to three times daily for 15 minutes for at least three months. In three of these studies (Onozawa 2001; Oswalt 2007; O'Higgins 2008), mothers were taught infant massage as part of a weekly group-based session. In the Onozawa 2001 study, mothers attended weekly group-based sessions for 70 minutes over the course of five weeks. The class leaders were trained by an International Association of Infant Massage teacher (IAIM) who aimed to encourage parents to observe and respond to their infant's cues and adjust their touch accordingly. In O'Higgins 2008, mothers attended weekly one-hour long classes over six weeks. In Oswalt 2009, the massage sessions were delivered as part of a parent training class where mothers were trained in massage. Infants were massaged for approximately 30 minutes daily for two months.

It was unclear from the Maimaiti 2007 study for how long the intervention was delivered, but it appears to have extended beyond the immediate postnatal period as the parents were instructed to continue massage once they had left the hospital with their infants. In two studies the intervention was delivered over a longer term, with the massage being performed one or two times a day from birth to six months of age (and continuing after six months of age). Massage lasted for 15 minutes each session in Jing 2007 and a minimum of 10 minutes massage daily over 16 weeks in Elliott 2002.

### Types of massage

It was clear from the small number of studies where information was provided about the massage technique, that the intensity or the amount of pressure applied during the massage varied from study to study. In Arikani 2008, massage was described as 'chiropractic spinal manipulation', but was derived from the method of Huhtala 2000, which is a gentle type of stroking massage. Jing 2007 used a massage and motion training method promoted by Johnson and Johnson (Johnson 2011), which comprises a gentle full-body massage including 'pedaling' motions of the legs, and opening and closing the arms. Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Zhu 2010 also used the Johnson and Johnson method. In Koniak-Griffin 1988, infants were massaged using a six-step sequential, cephalocaudal progression of stroking and gentle massage of the ventral and dorsal surfaces of the infant's body. In Kim 2003 researchers were trained to stroke each part of the infant's body in sequence and the process, intensity and pace of the intervention was agreed and reliability maintained at 96% during the course of the study. In Cigales 1997 the infants were massaged only on one occasion prior to an habituation task and this massage is described as deep but gentle massage of the whole body. Argawal 2000 used a standardised regimen based on traditional Swedish Massage practices. The mothers were given instructions and training for uniformity of massage strokes in terms of technique (force and direction) and time spent massaging individual body parts. Jump 1998; Elliott 2002; Oswalt 2007 and Onozawa 2001 do not describe the amount of pressure used, although a detailed description of the massage technique was given in Onozawa 2001, which

includes a full body massage using slow rhythmic strokes. Field 1996 gives a detailed description of each massage stroke and ensured that the researchers applied the correct intensity and pressure. Narenji 2008 described a full body massage using circular smooth movements (avoiding the eye and genital areas), but it is unclear how much pressure was used.

It was not possible to obtain this information from many of the studies reported in Chinese because the reports are short and we were unable to obtain further information from the trial investigators. A variety of techniques and amounts of pressure were used. For example, Ke 2001 describes how an additional method of kneading the back was added to the traditional massage method, and Maimaiti 2007 gives a detailed description of a full body massage using gentle pressing and sliding movements.

A small number of studies identified the importance of parent-infant communication during the delivery of the infant massage. O'Higgins 2008 stated that the emphasis was on paying attention to infant cues such that different massage strokes and amounts of massage could be tailored to each mother-infant pair. Onozawa 2001 and Oswalt 2007 also described how parents were taught to recognise and be sensitive to infant cues before commencing massage and throughout the massage as well. White-Traut 2009 used moderate pressure massage strokes and monitored the infants' behavioural responses prior to applying ATVV components of the massage. Cheng 2004 also encouraged parents to respond appropriately to infant cues, and to stop the massage if the baby cried or was tense.

### Outcomes

#### Types of outcome measures

Six studies (Koniak-Griffin 1988; Field 1996; Argawal 2000; Kim 2003; Jing 2007; Narenji 2008), assessed the impact of massage on physical outcomes including height, weight and physical growth. Field 1996 also measured formula intake.

The other studies (all conducted in China) that measured physical outcomes, assessed the impact of massage on growth (Wang 1999; Ke 2001; Zhai 2001; Shi 2002; Cheng 2004; Sun 2004; Liu CL 2005; Liu DY 2005; Lu 2005; Na 2005; Shao 2005), sleep (Sun 2004; Xua 2004; Liu DY 2005), bilirubin levels (Sun 2004; Lu 2005), sleep and crying (Cheng 2004; Xua 2004) and on incidence of common illnesses (Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months).

Argawal 2000 investigated the effect of different massage oils on physical growth and on physiological changes in blood flow and vessel diameter. Field 1996 measured levels of cortisol, epinephrine, norepinephrine and serotonin before and after massage, and Ferber 2002 measured 6-sulphatoxyymelatonin in urine. White-Traut 2009 measured salivary cortisol.

Five studies assessed the impact of massage on the mother-infant relationship (Jump 1998; Onozawa 2001; O'Higgins 2008).

[Elliott 2002](#) and [Koniak-Griffin 1988](#) reported mother and child interactions using the Nursing Child Feeding Assessment Scale (NCAFS), and the Nursing Child Teaching Assessment Scales (NCATS), and the Murray ratings scales. [O'Higgins 2008](#) also explored attachment patterns using the Strange Situation procedure. [Jump 1998](#) and [Onozawa 2001](#) both reported parenting stress using the Child Domain of the Parenting Stress Index.

Other outcomes included infant temperament measured using the Colorado Child Temperament Inventory, Infant Behaviour Questionnaire and the Revised Infant Temperament Questionnaire ([Koniak-Griffin 1988](#); [Field 1996](#); [Jump 1998](#); [Elliott 2002](#)); maternal perceptions of child temperament using the Infant care Questionnaire (ICQ) were reported in [O'Higgins 2008](#), and infant development using the Bayley psychomotor and mental development indices (PDI and MDI) ([Koniak-Griffin 1988](#)). [Jing 2007](#) reported infant mental development using the Gessel Development Quotient.

Several studies evaluated the effects of massage on sleep using a range of measures ([Argawal 2000](#); [Ferber 2002](#); [Narenji 2008](#)). [Ferber 2002](#) also measured activity patterns. [Elliott 2002](#) and [Arikan 2008](#) reported the impact of massage on crying or fussing using the number of hours per day spent crying or fussing. [Field 1996](#) and [White-Traut 2009](#) also reported infant behavioural state after massage using the methods described by Thoman ([Thoman 1981](#); [Thoman 1987](#)).

Cognitive outcomes such as habituation were measured by [Cigales 1997](#), and distractibility in response to a brightly coloured toy was measured by [O'Higgins 2008](#).

Six further studies reported mental and cognitive/developmental outcomes: Bayley Mental Development Index (MDI) and Psychomotor Development Index (PDI) in [Liu C 2001 0 to 2 months](#) and [Liu C 2001 3 to 6 months](#); the Capital Institute of Children 0 to 3 Years Old Mental Checklist IQ Formula (China) in [Wang 2001](#); movement, sight and auditory tracking in [Maimaiti 2007](#); and MDI and PDI from the Levin Scales, adapted by the China Institute of Psychology and Child Development Quotient were used in [Zhu 2010](#). [Jing 2007](#) reported scores from the Gessel Development Quotient.

### Timing of outcome measurement

Outcomes were assessed immediately post-intervention (within four weeks of the end of the intervention unless otherwise stated in the analyses). For example, [White-Traut 2009](#) assessed salivary cortisol immediately after the cessation of massage and again 10 minutes later.

Follow-up outcomes were reported for weight in [Koniak-Griffin 1988](#) (at eight months) and in [Kim 2003](#) and [Jing 2007](#) (at six months); for length ([Kim 2003](#); [Jing 2007](#) at six months) and for head circumference ([Kim 2003](#) at six months). [Xua 2004](#) provided three- and six-month follow-up assessments of crying and sleep. One-year follow-up was provided for parent-infant interactions ([O'Higgins 2008](#)) and mental development ([Jing 2007](#)). Eight-month and 24-month follow-up of mental and psychomotor development was provided in one study ([Koniak-Griffin 1988](#)).

### Excluded studies

We excluded 26 studies. Eleven studies were not randomised ([Ineson 1995](#); [Pardew 1996](#); [Peláez-Nogueras 1997](#); [Fernandez 1998](#); [Clarke 2000](#); [Darmstadt 2002a](#); [Li 2002](#); [Fogaça 2005](#); [Lee 2006](#); [Yilmaz 2009](#); [Serrano 2010](#)). In six studies ([Stack 1990](#); [Peláez-Nogueras 1996](#); [Peláez-Nogueras 1997b](#); [Huhtala 2000](#); [Zhu 2000](#); [Field 2004](#)), the control group was inappropriate (there was no 'no treatment' control group). One study was excluded due to the use of an ineligible intervention ([Field 2000b](#)). Five studies were excluded because of ineligible populations: HIV-exposed ([Oswalt 2009](#)); lower gestational age and birthweight than normal ([Scafidi 1996](#)); population outside the age range eligibility criterion ([Cullen 2000](#); [Jump 2006](#)), or the study involved the use of animals ([Zhu 2010](#)). We excluded two studies ([Park 2006](#); [Im 2007](#)) because they examined the use of massage as pain relief after routine heel needlestick tests. [Jing L 2007](#) examined the use of a multimodal intervention comprising massage alongside the use of an educational toy, and it was not possible to extract the effects of the infant massage alone.

Full details can be found in the [Characteristics of excluded studies](#) table.

### Risk of bias in included studies

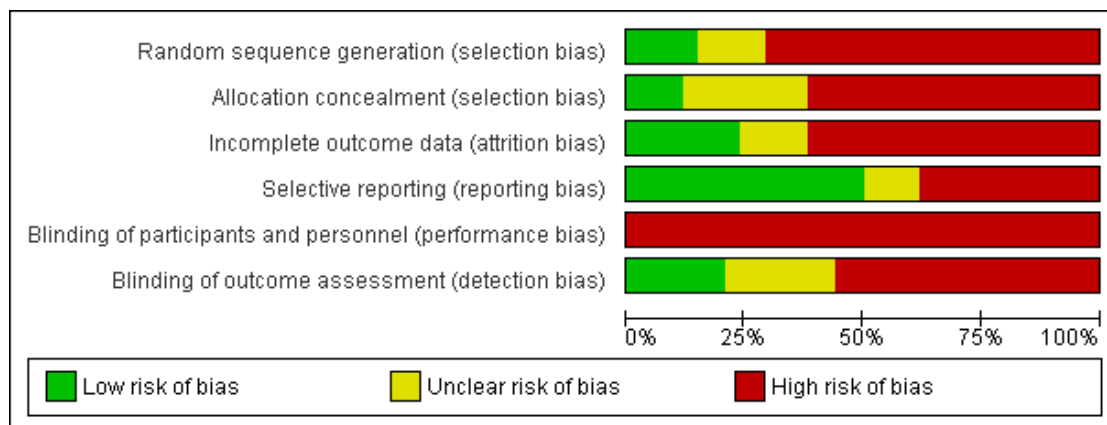
A summary of the risk of bias assessments across the 34 included studies is provided in [Figure 1](#) and [Figure 2](#).

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

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)
Argawal 2000	●	●	●	●	●	●
Arikan 2008	?	?	●	●	●	●
Cheng 2004	●	●	●	●	●	?
Cigales 1997	?	?	?	●	●	?
Duan 2002	●	●	●	●	●	●
Elliott 2002	●	●	●	●	●	●
Ferber 2002	?	?	●	●	●	?
Field 1996	●	?	●	●	●	?
Jing 2007	●	●	●	●	●	?
Jump 1998	●	?	?	?	●	●
Ke 2001	●	●	●	●	●	●
Kim 2003	●	?	?	●	●	?
Koniak-Griffin 1988	?	?	?	?	●	●
Liu C 2001 0 to 2 months	●	●	●	●	●	●
Liu C 2001 3 to 6 months	●	●	●	●	●	●
Liu CL 2005	●	●	●	●	●	●
Liu DY 2005	●	●	●	●	●	●
Lu 2005	●	●	●	●	●	●
Maimaiti 2007	●	●	●	●	●	●
Na 2005	●	●	●	●	●	●
Narenji 2008	●	●	●	●	●	●
O'Higgins 2008	●	?	●	●	●	●
Onozawa 2001	?	?	?	●	●	?
Oswalt 2007	●	●	●	?	●	●
Shao 2005	●	●	●	●	●	●
Shi 2002	●	●	●	●	●	●
Sun 2004	●	●	●	●	●	●
Wang 1999	●	●	●	●	●	●
Wang 2001	●	●	●	●	●	?
White-Traut 2009	●	●	●	●	●	●
Xua 2004	●	●	●	●	●	●
Ye 2004	●	●	●	●	●	●
Zhai 2001	●	●	●	?	●	●
Zhu 2010	●	●	●	●	●	●



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



## Allocation

### Randomisation

Fifteen studies were judged as high risk of bias because they were described as randomised but the study report provides insufficient details to be certain that the study was in fact randomised (Wang 1999; Ke 2001; Wang 2001; Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Elliott 2002; Shi 2002; Cheng 2004; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Liu DY 2005; Na 2005; Maimaiti 2007). We were unable to obtain any further details about the design of the study from the trial investigators to clarify this matter.

In five studies, insufficient details were provided about the method of randomisation to make a judgement about risk of bias and these were rated as unclear (Koniak-Griffin 1988; Cigales 1997; Onozawa 2001; Ferber 2002; Arikan 2008).

Nine studies were judged to be at high risk of bias because they used quasi-randomisation methods (Field 1996; Jump 1998; Zhai 2001; Kim 2003; Lu 2005; Shao 2005; O'Higgins 2008; White-Traut 2009; Zhu 2010). Of the quasi-randomised studies, two studies (Jump 1998; Kim 2003) used the flip of a coin to assign the first infant, and the remaining infants were alternately allocated to the intervention or control group. Lu 2005 and Zhu 2010 randomised according to the sequence of birth dates; Shao 2005 by sequence of birth time, and Zhai 2001 by odd or even hospital admission number. O'Higgins 2008 randomised according to availability of the intervention, using a prospective block-

controlled randomised design; mothers were contacted and invited to take part in either the massage group or the support group depending on which arm was recruiting at that given time point. White-Traut 2009 used a random number start in a table, then alternate allocation.

Only five studies were judged as low risk of bias in terms of the randomisation methods employed. These studies specified details of randomisation either in the study report or in further information obtained from the study investigator (Argawal 2000; Elliott 2002; Jing 2007; Oswalt 2007; Narenji 2008).

### Allocation concealment

Four studies described the method of allocation concealment. In Elliott 2002, a research associate who was not involved in the study assigned participants. Oswalt 2007; Narenji 2008; and White-Traut 2009 used sealed envelopes to conceal the allocation. Nine studies did not specify the method of allocation concealment and were judged as unclear risk of bias (Koniak-Griffin 1988; Field 1996; Cigales 1997; Jump 1998; Onozawa 2001; Ferber 2002; Kim 2003; Arikan 2008; O'Higgins 2008). The remaining studies did not apparently employ allocation concealment as there were no details in the study report, and as we were unable to obtain further details from the study investigator, we therefore judged these studies to be at high risk of bias.

### Blinding

### Blinding of participants and personnel

Blinding of the facilitators or parents who provided the infant massage intervention was not possible in the included studies due to the nature of the intervention, although in the [Field 1996](#) study nursery teachers and parents were unaware of the infants' allocation.

### Blinding of outcome assessors

Four studies ([Cigales 1997](#); [Koniak-Griffin 1988](#); [Elliott 2002](#); [Kim 2003](#)), used independent assessors who were blind to the intervention group. [Kim 2003](#) highlights the fact that despite precautions being taken to keep the orphanage staff blind to group assignment (staff members were out of the room during the intervention period), the staff may have become aware of the group assignment. In [Onozawa 2001](#), the assessment of mother-infant interaction scores was completed by the researcher who was aware of the infants' allocation groups. However, 10 dyads were coded by an experienced independent rater who was blind to study group and the researcher's reliability ratings were checked against the blinded coder. Two groups of dimensions did not meet the reliability standards and these were eliminated from the study.

[Ferber 2002](#) reported that both the actigraph measurements and the 6-sulphatoxyymelatonin secretions were analysed separately but does not clarify whether the assessors were blind to the participant group. [Jump 1998](#) did not use independent assessors.

[Cheng 2004](#) stated that the study was 'blind' but no further details were given, [Wang 2001](#) describes blind outcome assessment using a birth to three years of age development checklist, but it was unclear who was blinded and how this was achieved.

In the remaining studies, blinding of outcome assessors was either not attempted or not described with no further details provided, and these studies were judged to be at high risk bias.

### Incomplete outcome data

Five studies reported no dropout or attrition and these studies were judged at low risk of bias ([Field 1996](#); [Argawal 2000](#); [Arikan 2008](#); [Narenji 2008](#); [White-Traut 2009](#)). [Argawal 2000](#) was strictly regulated with mothers attending weekly to have their massage techniques monitored and to return empty oil bottles before collecting their next week's supply of specific oils. [Field 1996](#) reported no dropout for 40 postnatally depressed mother-infant dyads because the infants were being cared for by teachers in a nursery school during the six-week study. There was no dropout in [Arikan 2008](#), possibly because this intervention lasted for only one week. In [White-Traut 2009](#), the brief nature of the intervention resulted in no dropouts although insufficient sample volumes were collected for salivary cortisol analysis from all of the infants. No dropouts or losses to follow-up occurred in [Narenji 2008](#), according to further information from the trial investigator and we assessed this as low risk of bias.

Of the remaining studies that reported some dropout, [Jump 1998](#) reported a 21% dropout rate. Mothers from both groups who left the study were less educated and had younger infants than those remaining in the study, although the groups were otherwise alike demographically. It is unclear if this relatively high level of dropout introduced a risk of bias into the study. Fifteen per cent of mothers dropped out of [Elliott 2002](#) - five withdrew because they no longer met the eligibility criteria (the infants required hospital care), one infant was stillborn, four left because of family issues and seven dropped out because they found the study too time-consuming. [Ferber 2002](#) reported a dropout rate of 20% with no significant differences between the two intervention and control groups. [Koniak-Griffin 1988](#) reported a dropout rate of 2% at four months and 7% at eight months, mainly due to families moving out of the area. [Cigales 1997](#) excluded 34% of infants from the investigation due to excessive crying or fussing (n = 12), falling asleep (n = 3), experimenter error (n = 4) and fatigue (n = 1), which may have biased the results. In [Onozawa 2001](#), a total of 35% of the sample dropped out because the time of the class was inconvenient (seven from the massage and two from the control group did not complete and a further two mothers in the massage group and one in the control group did not have interactions recorded because their infants were unsettled). Although the dropouts were not evenly distributed between the groups, the infants who started and did not complete the study were not significantly different demographically from those that completed. It is unclear if this high level of dropout may have posed a risk of bias to the findings of the study. In [O'Higgins 2008](#), 31% did not complete in the massage group and 40% did not complete in the support group, with no statistical differences between the groups (that is 31 in each group completed the study to the first outcome assessment time point); we judged that this posed a low risk of bias to the study.

Only four studies undertook follow-up ([Koniak-Griffin 1988](#); [Kim 2003](#); [Jing 2007](#); [O'Higgins 2008](#)). [Kim 2003](#) lost 22% of 58 orphaned infants at the six-month follow-up, due to adoption. The loss was evenly spread between the groups, impacting on the power but not introducing a greater risk of bias into the study. [Koniak-Griffin 1988](#) presented data for only 41 children at four-, eight- and 24-months, representing an attrition rate of 39%. This was due in the main to families moving out of the area. Communication with the author confirmed that in the follow-up study, data were shown at four- and eight-months only for those 41 infants who had completed the study at 24 months. In [O'Higgins 2008](#), follow-up measures were reported at one year: 24 in the massage group completed all follow-up assessments, compared with 16 in the control (support only) group (further details about dropout were provided by the investigator). In [Jing 2007](#), it is unclear how many infants were lost to follow-up at the six-month time point. Numbers lost to follow-up are provided in the published report for the post-intervention time point, but the investigators were unable to supply any further details, including

reasons for loss to follow-up.

Nineteen studies reported that the same number who were recruited to the study completed the intervention and assessments but dropout or loss to follow-up was not addressed in the study report (Wang 1999; Ke 2001; Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Wang 2001; Zhai 2001; Duan 2002; Shi 2002; Cheng 2004; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Liu DY 2005; Lu 2005; Na 2005; Shao 2005; Maimaiti 2007; Zhu 2010). As a result of the fact that no information was provided in the published reports about dropout or loss to follow-up, and no further information was available from the trial investigators, we judged these studies to be at high risk of bias.

### Selective reporting

Reporting bias was unclear in four studies (Koniak-Griffin 1988; Jump 1998; Zhai 2001; Oswalt 2007). In Jump 1998, only questionnaire results at 12 months are reported; in Koniak-Griffin 1988 although all three components of the Bayley scales of infant development were administered, only the MDI and PDI findings were reported. In Oswalt 2007, mothers were asked to complete a worksheet, but no worksheets were completed and returned. In Zhai 2001, all the pre-specified outcomes were reported but milk intake was also reported, therefore it is unclear if other outcomes were measured but not reported. We judged the risk of bias as unclear.

Thirteen studies either did not pre-specify outcomes or provided insufficient information about outcome measurements (Wang 1999; Ke 2001; Duan 2002; Shi 2002; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Liu DY 2005; Lu 2005; Na 2005; Shao 2005; Maimaiti 2007). We were unable to obtain clarification from the trial investigators and judged these studies as being at high risk of bias.

The remaining studies were judged to be at low risk of reporting bias.

### Other potential sources of bias

#### Intention-to-treat analysis

None of the included studies explicitly stated that they were conducted on an intention-to-treat basis.

#### Distribution of confounders

While the use of randomisation should in theory ensure that any possible confounders are equally distributed between the arms of the trial, small numbers of trial participants may result in an unequal distribution of confounding factors. It is therefore important that the distribution of known potential confounders is: a) compared between the different study groups at the outset or b) adjusted for at the analysis stage.

Fourteen studies (Koniak-Griffin 1988; Field 1996; Jump 1998; Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Onozawa 2001; Elliott 2002; Ferber 2002; Kim 2003; Oswalt 2007; Arian 2008; Narenji 2008; O'Higgins 2008; White-Traut 2009), provided a detailed description or an analysis of the distribution of baseline demographic factors.

Fourteen studies provided a limited assessment of only a few potential confounders. Jing 2007 provided baseline measurements of weight and length, but no other demographic details; Cheng 2004 and Duan 2002 provided baseline weight, length and head circumference; Wang 1999; Wang 2001; Sun 2004; Liu CL 2005; Liu DY 2005; Lu 2005; Shao 2005; Zhu 2010 provided APGAR score and baseline weight; Sun 2004 and Zhu 2010 provided APGAR, baseline weight and maternal age; Cigales 1997 assessed maternal age and ethnicity.

Seven studies (Argawal 2000; Ke 2001; Shi 2002; Xua 2004; Ye 2004; Na 2005; Maimaiti 2007) did not analyse the distribution of confounders.

The intervention and control groups did not differ significantly in terms of demographic details in any of the included studies.

### Effects of interventions

In the text below, numbers given are the total number of participants randomised. Where it has been possible to calculate an effect size, we have reported these with 95% confidence intervals (CIs). Where we calculated and reported effect sizes, a minus sign indicates that the results favour the intervention group. Where the calculated effect size is statistically significant ( $P < 0.05$ ), we state whether the result favours the intervention or control condition.

In terms of effect sizes, values  $> 0.70$  have been treated as large; those between  $0.40$  and  $0.70$  as moderate; values  $< 0.40$  and  $> 0.10$  have been treated as small, and values  $< 0.10$  have been treated as no evidence of effectiveness (Higgins 2009, section 12.6.2).

An  $I^2$  value for heterogeneity is only reported as substantial if it exceeds 50% or if the  $P$  value from the  $\text{Chi}^2$  test is  $< 0.05$ .

For the purpose of subgroup analysis, duration of the massage programmes was categorised as follows: brief: a single session; short-term: up to four weeks; medium-term: four to 12 weeks; and long-term: more than 12 weeks.

We have summarised results below under headings corresponding to the outcomes outlined in the section entitled **Types of outcome measures**. For each outcome, we have presented the results according to the timing of the outcome assessment.

Under each heading, results of sensitivity analyses are included where these were conducted.

The results are organised as follows.

Results of studies comparing massage versus control group

1. Physical health and growth outcomes
2. Mental health and development outcomes

For each outcome, we present subgroups by timing of outcome assessment and provide the results of meta-analyses where data from more than one study could be combined.

## Massage versus control group: physical health and growth outcomes

### Weight

### Post-intervention

#### *Meta-analysis*

A meta-analysis of 18 studies of 2271 participants in total provided data for analysis of weight gain immediately post-intervention, and showed a significant increase favouring the experimental (massage) group ([Analysis 1.1](#)) (mean difference (MD) -965.25 g; 95% CI -1360.52 to -569.98). Heterogeneity was substantial (100%), and a number of sensitivity analyses were conducted.

#### *Sensitivity analyses*

In the previous published version of the review, we conducted sensitivity analyses to investigate the impact of one large study of orphaned infants ([Kim 2003](#)) in terms of weight ([Analysis 1.1](#)) because this population may be clinically different from the other participants (that is, in terms of the type of delivery of general care and levels of nurturing). We repeated this analysis for reasons of consistency, but removal of this study at this update did not affect the statistical significance of the result (MD -975.96 g; 95% CI -1390.63 to -561.30), and heterogeneity remained substantial at 100%.

We explored reasons for heterogeneity in further sensitivity analyses. When only studies carried out in the West were included in the analysis ([Koniak-Griffin 1988](#); [Field 1996](#)), the result favoured neither the intervention nor the control (MD -127.10 g; 95% CI -575.14 to 320.93; [Analysis 1.1](#)) and no significant heterogeneity was observed ( $I^2 = 0\%$ ).

We performed an additional sensitivity analysis to explore selection bias due to inadequate randomisation. When we included only those studies that we rated as adequately randomised, the result for weight gain at post intervention (from [Argawal 2000](#); [Jing 2007](#); [Narenji 2008](#)) again favoured neither the intervention nor the control group (MD -203.55 g; 95% CI -443.37 to 36.26).

#### *Subgroup analyses for duration of intervention*

We conducted subgroup analyses to assess whether the duration of the intervention affected the outcome. No brief intervention studies contributed growth outcome data, and the result of this analysis showed results favouring the intervention for massage programmes

of all durations: short-term interventions, five studies of 443 participants ([Argawal 2000](#); [Shi 2002](#); [Kim 2003](#); [Na 2005](#); [Narenji 2008](#)) (MD -374.07 g; 95% CI -654.84 to -93.31; [Analysis 1.2](#)), heterogeneity was substantial ( $I^2 = 93\%$ ); medium-term interventions, 12 studies of 1648 participants ([Koniak-Griffin 1988](#); [Field 1996](#); [Wang 1999](#); [Ke 2001](#); [Wang 2001](#); [Duan 2002](#); [Cheng 2004](#); [Sun 2004](#); [Ye 2004](#); [Liu CL 2005](#); [Lu 2005](#); [Liu DY 2005](#)) (MD -1259.19 g; 95% CI -1807.80 to -710.58; [Analysis 1.2](#)), heterogeneity was substantial ( $I^2 = 100\%$ ), and long-term, one study ([Jing 2007](#)) of 180 participants (MD -500.00 g; 95% CI -811.25 to -188.75; [Analysis 1.2](#)).

### Follow-up

#### *Meta-analysis*

Three studies of 202 participants in total provided follow-up data ([Kim 2003](#); [Jing 2007](#) at six months; [Koniak-Griffin 1988](#) at eight months). The finding was statistically significant in favour of the intervention (MD -758.29 g; 95% CI -1364.67 to -151.90; [Analysis 1.1](#)), but heterogeneity was substantial ( $I^2 = 81\%$ ). This significant result was largely due to impact of one study ([Kim 2003](#)), the remaining two studies showing no evidence of effectiveness.

#### *Sensitivity analyses*

We conducted sensitivity analyses to investigate the impact of one large study of orphaned infants ([Kim 2003](#)) in terms of weight at follow-up because this population may be clinically different from the other participants (see above). Removal of this study from the meta-analysis of follow-up data did not affect the statistical significance of the result (MD -455.07 g; 95% CI -823.80 to -86.33), but heterogeneity was reduced ( $I^2 = 0\%$ ).

### Length

### Post-intervention

#### *Meta-analysis*

Eleven studies of 1683 participants in total measured infant length at post-intervention. The result was statistically significant, favouring the intervention (MD -1.30 cm; 95% CI -1.60 to -1.00; [Analysis 1.3](#)). Heterogeneity was again substantial ( $I^2 = 80\%$ ).

### *Sensitivity analyses*

A sensitivity analysis in which we included only those studies rated as methodologically adequate (that is, having a low risk of bias due to randomisation) (Argawal 2000; Jing 2007; Narenji 2008) was still significant, favouring the intervention (MD -0.65 cm; 95% CI -1.20 to -0.11; Analysis 1.3). Heterogeneity was reduced, but still substantial ( $I^2 = 58\%$ ), and no further sensitivity analyses based on location (for example, Western versus Eastern studies) was possible.

### *Subgroup analyses for duration of intervention*

No studies of brief interventions contributed growth outcome data. The results show that duration of intervention did not affect significance of the result (that is, favoured the intervention irrespective of duration) (Analysis 1.4). For short-term interventions, we included five studies of 443 participants (Argawal 2000; Shi 2002; Kim 2003; Na 2005; Narenji 2008) (MD -1.00 cm; 95% CI -1.54 to -0.47) and heterogeneity was substantial ( $I^2 = 70\%$ ); medium-term term-interventions involved five studies of 1060 participants (Ke 2001; Duan 2002; Cheng 2004; Liu DY 2005; Lu 2005) (MD -1.51 cm; 95% CI -1.76 to -1.27), with reduced but substantial heterogeneity ( $I^2 = 53\%$ ); and one study of a long-term intervention (Jing 2007) involving 180 participants (MD -1.13 cm; 95% CI -1.88 to -0.38; Analysis 1.4).

### **Follow-up**

#### *Meta-analysis*

Jing 2007 and Kim 2003 evaluated the effectiveness of massage on infant length. A meta-analysis comprising 161 participants at six months post-intervention found that the significant increase in the intervention group had not been maintained (MD -1.98 cm; 95% CI -4.69 to 0.72; Analysis 1.3). Heterogeneity was again substantial ( $I^2 = 87\%$ ).

### **Head circumference**

#### **Post-intervention**

#### *Meta-analysis*

Nine studies reported head circumference at post-intervention. A meta-analysis comprising 1423 participants produced a significant result favouring the intervention (MD -0.81 cm; 95% CI -1.18 to -0.45; Analysis 1.5). Heterogeneity was substantial ( $I^2 = 87\%$ ).

### *Sensitivity analyses*

We performed a sensitivity analysis in which we included only the two studies that we rated as being at low risk of selection bias (randomisation) (Argawal 2000; Narenji 2008). The result favoured neither the intervention nor the control (MD -0.07 cm; 95% CI -0.27 to 0.12;  $I^2 = 0\%$ ; Analysis 1.5). No further analyses based on location were possible.

### *Subgroup analyses for duration of intervention*

No studies provided growth outcome data following brief or long-term infant massage. The results of the remaining two subgroup analyses are presented in Analysis 1.6. For short-term interventions, four studies contributed 363 participants (Argawal 2000; Kim 2003; Na 2005; Narenji 2008) (MD -0.70 cm; 95% CI -1.45 to 0.05) with no evidence of effectiveness, and substantial heterogeneity ( $I^2 = 89\%$ ); for medium-term duration interventions, five studies contributed 1060 participants (Ke 2001; Duan 2002; Cheng 2004; Liu DY 2005; Lu 2005) and the result favoured the intervention (MD -0.90 cm; 95% CI -1.16 to -0.64), and heterogeneity was again substantial ( $I^2 = 58\%$ ), but no sensitivity analysis was possible.

### **Follow-up**

Two studies (Kim 2003; Zhu 2010) reported growth outcome data at six-month follow-up with the result favouring the intervention (MD -2.19 cm; 95% CI -3.88 to -0.49; Analysis 1.5). Heterogeneity was substantial ( $I^2 = 91\%$ ).

### **Mid-arm/mid-leg circumference**

#### **Post-intervention**

#### *Meta-analysis*

Two studies (Argawal 2000; Narenji 2008) evaluated the impact of infant massage on mid-arm (Analysis 1.7) and mid-leg (Analysis 1.8) circumference at post-intervention. The meta-analyses, each comprising 225 participants, showed statistically significant results favouring the intervention group (MD -0.47 cm; 95% CI -0.80 to -0.13) for the arm measurement (Analysis 1.7); and for the leg measurement (MD -0.31 cm; 95% CI -0.49 to -0.13). Heterogeneity was substantial for the mid-arm measurement ( $I^2 = 80\%$ ), but low ( $I^2 = 0\%$ ) for the mid-leg measurement, but not sensitivity analysis was possible.

## Abdominal and chest circumference

### Post-intervention

#### *Single study results*

Only [Narenji 2008](#) measured abdominal and chest circumference at post-intervention and therefore, no meta-analysis was possible. There was a statistically significant result for this single study, favouring massage for both abdominal circumference (MD -0.75 cm; 95% CI -1.09 to -0.41; [Analysis 1.9](#)) and chest circumference (MD -0.88 cm; 95% CI -1.22 to -0.54; [Analysis 1.10](#)).

### Other study results

The following studies provided means and significance levels only, and these data could not therefore be entered into a meta-analysis. Data from a six-month vertical survey of the growth of all ( $n = 310$ ) the infant participants over zero to six months in two studies ([Liu C 2001 0 to 2 months](#); [Liu C 2001 3 to 6 months](#)) showed significant differences in the weight and the chest circumference of the infants who received the massage. Height and head circumference were not significantly different (study results summarised in [Table 1](#)). [Maimaiti 2007](#) provided post-intervention assessments for weight, length and head circumference and found significant differences between massage and control groups ( $P > 0.05$ ), [Table 1](#). Two further studies provided means and significance levels only ([Zhai 2001](#); [Shao 2005](#)). The results for both studies indicated significant findings favouring the intervention groups.

## Hormones

### Post-intervention

#### *Meta-analysis*

Two studies ([White-Traut 2009](#); [Field 1996](#)) measured salivary cortisol levels using units of  $\mu\text{g/dL}$  ([White-Traut 2009](#)) and  $\text{ng/mL}$  ([Field 1996](#)) at 10 and 20 minutes respectively, after the completion of the massage interventions. Although [White-Traut 2009](#) reported that cortisol levels measured at 10 minutes after the intervention had declined, meta-analysis of 54 participants from [White-Traut 2009](#) and [Field 1996](#) showed no significant difference between groups (SMD -0.24; 95% CI -0.77 to 0.30; [Analysis 1.11](#)).

#### *Single study results*

A number of other studies reported findings for hormones but these could not be pooled in a meta-analysis. [White-Traut 2009](#) reported that salivary cortisol levels ( $\mu\text{g/dL}$ ) were higher immediately after a single session of the intervention in the massage group. This was not statistically significant in our analyses (standardised mean difference (SMD) 0.46; 95% CI -0.45 to 1.38; [Analysis 1.11](#)).

[Field 1996](#) measured urinary cortisol ( $\text{ng/mL}$ ) using radioimmunoassay on day 12 of the intervention, and this was significantly lower in the massage group (SMD -0.80; 95% CI -1.45 to -0.15; [Analysis 1.11](#)).

[Field 1996](#) measured norepinephrine, epinephrine and serotonin in urine samples, which were frozen and sent for high-pressure liquid chromatography assays with electrochemical detection. Results showed significant improvements for the treatment group including reduced levels of norepinephrine (MD -60.30; 95% CI -111.88 to -8.72; [Analysis 1.12](#)) and epinephrine (MD -13.00; 95% CI -20.08 to -5.92; [Analysis 1.13](#)). A non-significant result was reported for levels of serotonin (MD -295.50; 95% CI -705.25 to 114.25; [Analysis 1.14](#)).

[Ferber 2002](#) evaluated the effect of massage therapy on the nocturnal secretion of 6-sulphatoxymelatonin in urine ( $\text{ng}$ ). The results indicated significantly higher levels in the massaged group (MD -523.03; 95% CI -664.51 to -381.55; [Analysis 1.15](#)).

## Biochemical markers

### Post-intervention

#### *Meta-analysis for bilirubin*

Two studies ([Sun 2004](#); [Lu 2005](#)) with a sample of 410 (205 intervention and 205 control) measured bilirubin ( $\text{mmol/L}$ ) seven days after birth and found significantly lower levels in the massaged infants (MD -38.11  $\text{mmol/L}$ ; 95% CI -50.61 to -25.61; [Analysis 1.16](#)). Heterogeneity was substantial ( $I^2 = 52\%$ ). No sensitivity analysis was possible.

### Activity cycle

### Post-intervention

#### *Single study results*

At eight-weeks postnatal, [Ferber 2002](#) observed peak activity during the time period 3 am to 7 am in the massaged group treatment group compared with 11 pm to 3 am in the control group. A secondary peak of activity was observed in the treated children between 3 pm and 7 pm while in the control group a secondary peak occurred between 11 am to 3 pm. The interaction between treatment and timing of peak activity was statistically significant ( $P = 0.042$ ). This suggests a delay in peak activity in massaged infants, and that the treated infants achieved a more favourable adjustment of their rest-activity cycle ([Ferber 2002](#)). No significant differences were found between groups in total movement. No differences were found for measurements performed one-day before and one-day after the intervention and at six-weeks of age (study results, no analysis possible).

### **Behaviours including crying and fussing time and sleep/wake behaviours**

#### **Post-intervention**

#### *Meta-analysis*

A meta-analysis of 341 participants in total, from four studies ([Elliott 2002](#); [Cheng 2004](#); [Xua 2004](#); [Arikan 2008](#)), showed no significant difference in the number of hours per day spent crying or fussing (MD -0.36; 95% CI -0.52 to -0.19; [Analysis 1.17](#)).

#### *Single study results*

[Xua 2004](#) reported crying frequency, that is the number of episodes of crying. Infants in the massage group cried less often than the control group at all time points and this was statistically significant at all time points ([Analysis 1.18](#)), including post-intervention (MD -0.34; 95% CI -0.56 to -0.12).

[Field 1996](#) assessed sleep/wake behaviours using an adaptation of the system of sleep recording developed by [Thoman 1981](#). Significantly less crying (MD -8.20; 95% CI -12.24 to -4.16), more increased active awake behaviour (MD -15.00; 95% CI -22.29 to -7.71) and significantly more time in an inactive alert state (MD -12.70; 95% CI -19.38 to -6.02) was observed in the massage group. Measures of quiet sleep (MD -6.30; 95% CI -20.16 to 7.56) and movement (MD -12.60; 95% CI -27.59 to 2.39) favoured neither the intervention nor the control group. There was also no significant difference between massage and control groups in the amount of drowsiness (MD 2.00; 95% CI -0.19 to 4.19; [Analysis 1.19](#)).

[White-Traut 2009](#) assessed behavioural state ([Thoman 1987](#)) immediately post-intervention (after a single instance of massage). There were no significant differences between the groups in the

number of infants asleep (risk ratio (RR) 1.04; 95% CI 0.55 to 1.96), awake (RR 0.78; 95% CI 0.27 to 2.23) or crying (RR 1.94; 95% CI 0.09 to 43.50) ([Analysis 1.20](#)).

#### **Follow-up**

#### *Single study results*

[Xua 2004](#) recorded the number of hours per day spent crying or fussing at follow-up, in [Analysis 1.17](#). The result was significant (favouring the intervention) at the three-month follow-up: (MD -0.21 95% CI -0.40 to -0.02); and at the six-month follow-up (MD -0.15 95% CI -0.29 to -0.01).

[Xua 2004](#) also reported crying frequency at follow-up. Infants in the massage group cried significantly less often than the control group at the three-month follow-up (MD -0.19; 95% CI -0.36 to -0.02); and the six-month follow-up (MD -0.18; 95% CI -0.35 to -0.01) ([Analysis 1.18](#)).

#### **Sleep habits**

#### **Post-intervention**

#### *Meta-analysis*

A meta-analysis of data from four studies ([Sun 2004](#); [Xua 2004](#); [Liu DY 2005](#); [Narenji 2008](#)) ( $n = 634$  participants), found a significant difference in 24-hour sleep duration, favouring the massage group (MD -0.91 hr; 95% CI -1.51 to -0.30; [Analysis 1.21](#)). Heterogeneity was substantial ( $I^2 = 94\%$ ).

For mean increase in hours of sleep over a 24-hour period, a meta-analysis of participant data from two studies ([Argawal 2000](#); [Narenji 2008](#)) ( $n = 225$ ) favoured neither the intervention nor the control (SMD -1.47; 95% CI -4.43 to 1.49; [Analysis 1.22](#)).

[Argawal 2000](#) and [Narenji 2008](#) contributed 225 participants to a meta-analysis of mean increase in duration of night sleep. The results were not statistically different (SMD -1.28; 95% CI -3.66 to 1.10; [Analysis 1.23](#)). Heterogeneity was substantial ( $I^2 = 98\%$ ), but no sensitivity analysis was possible.

#### *Single study results*

Measurements of sleep were reported using a variety of other measures but few were sufficiently similar to permit pooling of data in a meta-analysis.

Argawal 2000 reported the increase in duration of daytime sleep, and the result favoured neither the intervention nor the control (MD 0.10 hr; 95% CI -0.21 to 0.41; Analysis 1.24).

Argawal 2000 recorded a mean increase in the duration of the first morning sleep after massage, favouring the intervention (MD -1.52; 95% CI -1.69 to -1.35; Analysis 1.25).

Narenji 2008, observed a significant increase in favour of the intervention in the total number of hours sleep per night (MD -0.70 hr/night; 95% CI -1.00 to -0.40; Analysis 1.26).

Argawal 2000 assessed the number of naps (short periods of sleep). There were approximately one fewer naps for both groups (0.7 compare with 0.5) respectively (MD -0.22; 95% CI -0.55 to 0.11; Analysis 1.27), although this difference was not statistically significant. There was no statistical difference between intervention or control for the number of naps during the day or at night (Analysis 1.28 and Analysis 1.29, respectively).

Xua 2004 reported night wake frequency (the number of times the infant woke per night). Infants woke significantly less often in the massage group than in the control group at post-intervention (MD -0.48; 95% CI -0.81 to -0.15; Analysis 1.30). Xua 2004 also reported the duration of night wake periods: infants were awake at night for significantly less time in the massage group compared with the control group (the control group was awake at night on average for 16 minutes longer at post-intervention (Analysis 1.31). Sleep habits at post-intervention were also reported in two studies (Liu C 2001 0 to 2 months and Liu C 2001 3 to 6 months.), and were categorised as ('good', 'medium', and 'not good') but means and standard deviations were not provided and meta-analysis was not therefore possible. The results showed significantly more 'good' sleepers in newborn to two-month infants ( $X^2 = 15.353$ ;  $P = 0.0000$ ; Table 2), but not in the 3 to 6 month old infants.

## Follow-up

### Single study results

No meta-analysis was possible for 24-hour sleep duration at three- or six-month follow-up because only one study reported data for these time points (Xua 2004). At three months, the result significantly favoured the intervention (infants slept longer over 24 hours) (SMD -1.30 95% CI -1.81 to -0.79; Analysis 1.21) but by six-month follow-up there was no difference between the intervention and control group infants (Analysis 1.21).

Xua 2004 reported night wake frequency (that is, the number of times the infant woke per night). Infants woke significantly less often in the massage group than in the control group at post-intervention (see above) and at the three-month (MD -0.38; 95%

CI -0.63 to -0.13; Analysis 1.30) and six-month follow-up (MD -0.35; 95% CI -0.56 to -0.14; Analysis 1.30).

Xua 2004 also reported the duration of night wake periods: infants were awake at night for significantly less time in the massage group compared with the control group (the control group was awake at night on average for longer at post-intervention (see above), and this trend continued at follow-up: the control infants were awake 10 minutes longer than the massaged infants at the three-month, and 15 minutes longer at the six-month follow-up; Analysis 1.31).

## Blood flow

### Post-intervention

#### Single study results

Argawal 2000 assessed the impact of infant massage on blood velocity, vessel diameter and blood flow after four weeks of massage. The results were not significant for blood velocity (MD -0.98 cm; 95% CI -6.65 to 4.69; Analysis 1.32), but there was a significant difference for vessel diameter favouring the control group (MD 0.02 cm; 95% CI 0.01 to 0.03; Analysis 1.32); and for blood flow but favouring the massage group (MD -0.54 cm; 95% CI -1.03 to -0.05; Analysis 1.32).

## Formula intake

### Post-intervention

#### Single study results

Field 1996 measured the impact of massage on formula intake. No units of measurement were provided in the published paper, but we have assumed that US fl. oz was used, and have converted the values to mL. The results indicated a significantly higher intake in the control group of just over 70 mL of formula (MD 70.97 mL; 95% CI 6.16 to 135.78; Analysis 1.33).

## Number of illnesses and clinic visits

### Post-intervention

### Meta-analysis



A meta-analysis comprising 310 participants from two studies (Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months), showed that fewer infants suffered from diarrhoea in the massage group (RR 0.39; 95% CI 0.20 to 0.76), Analysis 1.34. There was no heterogeneity between the studies.

There were no significant differences between intervention and control groups in the number of episodes of upper respiratory tract infection (URTI) or anaemia for either Liu C 2001 0 to 2 months or Liu C 2001 3 to 6 months.

## Follow-up

### Single study results

At the six-month follow-up, there was a significant reduction in the number of illnesses (MD -8.82; 95% CI -10.62 to -7.02;  $P = 0.00001$ ), and clinic visits (MD -5.98; 95% CI -7.07 to -4.89; Analysis 1.35) for intervention group orphanage infants compared with control group orphanage infants (Kim 2003).

## Massage versus control group: mental health and developmental outcomes

### Infant temperament

#### Post-intervention

### Meta-analysis

A meta-analysis of activity sub-scale scores from the Colorado Child Temperament Inventory, Infant Behaviour Questionnaire, and the Revised Infant Temperament Questionnaire post-intervention comprising data from 121 participants from three studies (Koniak-Griffin 1988; Field 1996; Jump 1998) (SMD 0.39; 95% CI -0.34 to 1.13; Analysis 2.1), showed no significant differences. Heterogeneity between the studies was substantial ( $I^2 = 75\%$ ), but no sensitivity or subgroup analyses were possible. There were also no significant differences for 'persistence' (data from 81 participants from Field 1996 and Koniak-Griffin 1988) or 'soothability' (80 participants from Field 1996; Jump 1998) (Analysis 2.1).

### Single study results

Field 1996 measured aspects of temperament using the Colorado Child Temperament Inventory (CCTI (Rowe 1977)). There was

no significant difference for activity, emotionality, sociability, persistence, or food adaptation (Analysis 2.2). Infants in the massage group were, however, statistically more likely to be soothable (soothability) (Analysis 2.2) (MD -2.90; 95% CI -5.71 to -0.09). Jump 1998 measured a range of aspects of infant temperament using the Infant Behaviour Questionnaire (IBQ). There were no differences between intervention and control groups for duration of orienting (MD 0.00; 95% CI -0.82 to 0.82); distress to limitations (MD -0.08; 95% CI -0.49 to 0.33); soothability (MD 0.03; 95% CI -0.59 to 0.65); fear (MD -0.06; 95% CI -0.63 to 0.51); or amount of smiling (MD 0.30; 95% CI -0.14 to 0.74). Infant activity level (MD 0.56; 95% CI 0.08 to 1.04; Analysis 2.3) significantly favoured the control group.

Koniak-Griffin 1988 used the Revised Infant Temperament Questionnaire (RITQ Carey) post-intervention. For each of the nine categories, a higher score (above the mean) generally denotes a trait that is deemed more negative and is indicative of a baby that is difficult or high-spirited. Lower scores (below the mean) are viewed as being more positive and indicative of an easy-to-parent baby. No significant differences were seen for eight of the nine measures: rhythmicity, approach, adaptability, intensity, mood, persistence, distractibility or threshold. Activity scores were significantly different and favoured the control group (MD 0.41; 95% CI 0.11 to 0.71; Analysis 2.4).

Elliott 2002 measured temperament using the nine scales comprising the Early Infant Temperament Questionnaire (Medoff-Cooper 1993), but did not provide adequate data to calculate effect sizes. She reported no significant group differences for any of the following: activity, rhythmicity, approach, adaptability, mood, persistence, distractibility, intensity or threshold.

## Follow-up

### Single study results

Koniak-Griffin 1988 found significant differences favouring the control group using the Revised Infant Temperament Questionnaire (RITQ Carey) at eight-month follow-up (Analysis 2.5), for rhythmicity (MD 0.80; 95% CI 0.12 to 1.48); approach (MD 0.88; 95% CI 0.25 to 1.51); adaptability (MD 0.69; 95% CI 0.01 to 1.37); intensity (MD 0.39; 95% CI 0.02 to 0.76); mood (MD 1.08; 95% CI 0.65 to 1.51); and distractibility (MD 0.72; 95% CI 0.32 to 1.12). There were no significant differences for activity, persistence or threshold.

### Infant Care Questionnaire

## Post-intervention and follow-up

### *Single study result*

O'Higgins 2008 assessed the impact of infant massage on infant characteristics using the Infant Care Questionnaire (ICQ). The results showed no significant differences at post-intervention (Analysis 2.6) or follow-up (Analysis 2.7) for any of the sub-scales (fussy/difficult; unadaptable; dull; unpredictable).

### **Infant attachment**

#### **Follow-up**

### *Single study results*

Jump 1998 measured infant attachment at one-year follow-up using the attachment Q-set (Waters 1985). The results for the whole sample indicated no significant effect on attachment security (MD -0.06; 95% CI -0.17 to 0.05; Analysis 2.8). (N.B. results reported in the study indicated a significant effect on infant attachment security in an 'as treated' analysis in which data for infants that had not complied with treatment were omitted).

### **Home environment**

#### **Follow-up**

### *Single study results*

Koniak-Griffin 1988 measured the impact of infant massage on the home environment at 24 months (based on a sub-sample of 49 infants in all four arms of the study, 12 in the experimental group and 13 in the control group) using the HOME Inventory (Bradley 1977). The findings showed no difference between groups (MD 0.34; 95% CI -1.92 to 2.60; Analysis 2.9).

### **Child behaviour**

#### **Follow-up**

### *Single study result*

Koniak-Griffin 1988 assessed the impact of infant massage on child behaviour at 24 months using the Eyberg Child Behaviour Inventory (ECBI) (Robinson 1980). The results showed a non-significant difference for intensity (MD 4.95; 95% CI -9.94 to

19.84; Analysis 2.10) and no effect for the problem domain (MD -0.19; 95% CI -3.26 to 2.88; Analysis 2.11).

### **Infant and mother-infant interactions**

#### **Post intervention**

### *Meta-analysis*

We combined data from three studies measuring mother and child interaction using either the total scores from the Nursing Child Teaching Assessment Scale (NCATS) (Koniak-Griffin 1988; Elliott 2002) or the Murray Global Rating Scale (Onozawa 2001; O'Higgins 2008) using data from 131 participants. The results favoured neither the intervention nor control group (SMD -0.26; 95% CI -1.01 to 0.48 ( $I^2 = 75%$ ); Analysis 2.12).

#### **Follow-up**

### *Meta-analysis*

A meta-analysis of follow-up results of 65 participants from Koniak-Griffin 1988 (at 24-month follow-up, based on a sub-sample of 15 out of 49 infants available for follow-up) and O'Higgins 2008 (at 12-month follow-up) was also not significant (favoured neither intervention nor control) (SMD -0.20; 95% CI -0.69 to 0.29 ( $I^2 = 0%$ ); Analysis 2.12).

#### **Post-intervention**

### *Single study results (Nursing Child Feeding Assessment Scale, NCAFS)*

Elliott 2002 found no significant differences between intervention and control groups using the Nursing Child Feeding Assessment Scale (NCAFS) (MD -2.10; 95% CI -1.96 to 6.16; Analysis 2.13).

#### **Follow-up**

### *Single study results (Nursing Child Teaching Assessment Scales, NCATS sub-scales)*

Koniak-Griffin 1988 measured the impact of infant massage on mother-infant interaction at 24-month follow-up (based on a sub-

sample of 49 infants) using the NCATS sub-scales. The results showed no significant improvement in mother-infant interaction for the Mother (SMD -0.18; 95% CI -0.96 to 0.61; [Analysis 2.14](#)) or Child sub-scales (SMD 0.35; 95% CI -0.44 to 1.14; [Analysis 2.15](#)).

## Post-intervention

### *Meta-analysis of Murray ratings sub-scales*

Two studies ([Onozawa 2001](#); [O'Higgins 2008](#)) reported the findings of video-recorded parent-infant interactions using a standardised coding schema ([Murray 1996](#)). All meta-analyses involved data from both studies for 84 participants for the following sub-scales: maternal sensitivity (warm to cold; intrusive to non-intrusive; remoteness); and infant interactions (attentive to inattentive; lively to inert; happy to distressed). The results showed no significant difference between groups: maternal sensitivity (warm to cold) (MD -0.34; 95% CI -1.07 to 0.40; [Analysis 2.16.1](#)) ( $I^2 = 91%$ ); intrusive to non-intrusive (MD -0.10; 95% CI -0.85 to 0.66; [Analysis 2.17.1](#);  $I^2 = 90%$ ); maternal remoteness (MD 0.08; 95% CI -0.32 to 0.48; [Analysis 2.18.1](#)); infant interactions - infant attentive to inattentive ([Analysis 2.19](#)); infant lively to inert ([Analysis 2.20](#)); or infant happy to distressed sub-scales ([Analysis 2.21](#)).

## Follow-up

### *Single study results (Murray ratings sub-scales)*

[O'Higgins 2008](#) found significant improvements favouring the intervention for only one aspect of maternal sensitivity at one-year follow-up (warm to cold) (MD -0.84; 95% CI -1.07 to -0.61; [Analysis 2.16](#)). There were no significant differences for maternal intrusive to non-intrusive measure, or maternal remoteness at one-year follow-up ([Analysis 2.17](#); [Analysis 2.18](#)).

There were no significant differences at one-year follow-up for Infant attentive to inattentive ([Analysis 2.19](#)); infant lively to inert ([Analysis 2.20](#)); or infant happy to distressed ([Analysis 2.21](#)) ([O'Higgins 2008](#)).

## Parenting stress (PSI)

### Post-intervention

### *Meta-analysis*

Two studies ([Jump 1998](#); [Oswalt 2007](#)) measured parenting stress using the child characteristics sub-scale of the PSI at post-intervention. The results of a meta-analysis of 55 participants at post-intervention, showed no significant difference between the two groups (MD -10.85; 95% CI -53.86 to 32.16; [Analysis 2.22](#)). Heterogeneity was substantial ( $I^2 = 91%$ ), but no sensitivity or subgroup analyses were possible.

## Psychomotor and mental development

### Post-intervention (PDI, psychomotor)

### *Meta-analysis*

Three studies ([Koniak-Griffin 1988](#); [Liu C 2001 0 to 2 months](#); [Liu C 2001 3 to 6 months](#)) evaluated the impact of infant massage on psychomotor development using the Bayley Scale of Infant Development ([Bayley 1969](#)). One further study ([Zhu 2010](#)) assessed psychomotor development using the Levin PDI (adapted by the China Institute of Psychology and Child Development Center). A meta-analysis of PDI scores from these four studies ([Koniak-Griffin 1988](#); [Liu C 2001 0 to 2 months](#); [Liu C 2001 3 to 6 months](#); [Zhu 2010](#)) with 466 participants in total, gave a significant result favouring the intervention group (SMD -0.35; 95% CI -0.54 to -0.15); [Analysis 2.23](#)).

### *Sensitivity and subgroup analyses*

A sensitivity analysis, using data from the single study conducted in the West ([Koniak-Griffin 1988](#)) indicated no difference between massage and control groups (SMD 0.00; 95% CI -0.61 to 0.62), [Analysis 2.23](#). We did not explore the potential effects of bias introduced by inadequate randomisation because all the studies were either at high risk ([Liu C 2001 0 to 2 months](#); [Liu C 2001 3 to 6 months](#); [Zhu 2010](#)) or were rated as unclear ([Koniak-Griffin 1988](#)). The intervention was of medium-term duration in all three studies.

### Follow-up (PDI, psychomotor)

### *Single study results*

[Koniak-Griffin 1988](#) measured the impact of infant massage on psychomotor at eight months using the Bayley PDI scales ([Bayley 1969](#)). The results show no effect for the PDI sub-scale (MD -0.78; 95% CI -11.89 to 10.33; [Analysis 2.24](#)).

At 24-month follow-up (Koniak-Griffin 1988) (based on a sub-sample of 41 infants), the results showed a non-significant difference in psychomotor development on the PDI sub-scale (MD -7.52; 95% CI -16.53 to 1.49; Analysis 2.24).

### Post-intervention (MDI, mental)

#### Meta-analysis

The Bayley Mental Development Index (MDI) scales were used to assess development in three studies (Koniak-Griffin 1988, Liu C 2001 0 to 2 months and Liu C 2001 3 to 6 months) and the Levin MDI (adapted by the China Institute of Psychology and Child Development Center) was used in one study (Zhu 2010). A meta-analysis of MDI scores from these four studies (Koniak-Griffin 1988; Liu C 2001 0 to 2 months; Liu C 2001 3 to 6 months; Zhu 2010) contributing 466 participants in total, gave a non-significant result (SMD -0.27; 95% CI -0.64 to 0.11; Analysis 2.25). Heterogeneity was substantial ( $I^2 = 69\%$ ).

#### Sensitivity and subgroup analyses

A sensitivity analysis, using data from the single study conducted in the West (Koniak-Griffin 1988), indicated a non-significant effect (SMD 0.38; 95% CI -0.23 to 1.00; Analysis 2.25). No further sensitivity or subgroup analyses were possible.

### Follow-up (MDI, mental)

#### Single study results

One study (Koniak-Griffin 1988) found a significant difference favouring the control group for mental development using the MDI sub-scale at eight-month follow-up (MD 22.85; 95% CI 4.26 to 41.44; Analysis 2.26). At 24-month follow-up (based on a sub-sample of 41 infants), the results were non-significant for mental development using the MDI scale (MD -8.59; 95% CI -18.80 to 1.62; Analysis 2.26).

### Other developmental measures

### Post-intervention

#### Meta-analysis

Jing 2007 and Wang 2001 utilised two different assessment scales (Gessel Developmental Quotient and Capital Institute Mental

checklist respectively), but four of the domains were sufficiently similar to combine in a meta-analysis of 237 participants (Analysis 2.27). The results were significant for gross motor (SMD -0.44; 95% CI -0.70 to -0.18); fine motor (SMD -0.61; 95% CI -0.87 to -0.35); and social behaviour (SMD -0.90; 95% CI -1.61 to -0.18); but not significant for language development (SMD -0.82; 95% CI -1.67 to 0.03). Both of these studies have been rated as being at high risk of bias.

#### Single study results

Jing 2007 measured aspects of development using the Gessel Developmental Quotient. For all five domains, there were significant differences favouring the intervention group (Analysis 2.28): adaptive behaviour (MD -7.07; 95% CI -9.75 to -4.39); gross motor (MD -3.97; 95% CI -6.99 to -0.95); fine motor (MD -6.89; 95% CI -10.18 to -3.60); language (MD -4.15; 95% CI -7.03 to -1.27); personal social behaviour (MD -6.41; 95% CI -9.65 to -3.17). There were significant gains in all aspects of development (gross motor, fine motor, cognitive, language, social behaviour) using the Capital Institute Mental Checklist (Wang 2001; Analysis 2.29), including a reported very large gain in IQ (MD -27.18; 95% CI -33.13 to -21.23) favouring the massage group.

One study that did not report means and standard deviations (Maimaiti 2007) and therefore could not be included in the numerical analyses presented here, assessed the extent to which infants could rise from a prone position, track objects visually (sight tracking), hearing (auditory tracking) and smile for the outcome assessors at post-intervention. All results are reported as being significant and favouring the massage group (Table 3).

### Follow-up

#### Single study results

Jing 2007 measured aspects of development at six-month follow-up using the Gessel Developmental Quotient. Four out of five domains (adaptive behaviour, fine motor, language and personal-social behaviour), showed significant differences favouring the intervention group (Analysis 2.30). Only the 'gross motor' domain failed to reach significance.

### Attachment (Strange Situation Procedure)

### Follow-up

### Single study results

O'Higgins 2008 examined the impact of infant massage on attachment using the Strange Situation Procedure at one-year follow-up. The finding showed no significant differences for any of the four sub-scales: secure (RR 0.82; 95% CI 0.50 to 1.34); avoidant (RR 1.39; 95% CI 0.14 to 14.07); resistant (RR 3.48; 95% CI 0.45 to 27.02); or disorganised (RR 0.70; 95% CI 0.16 to 3.02; Analysis 2.31).

### Distractibility

### Follow-up

### Single study results

O'Higgins 2008 examined distractibility in response to a brightly coloured toy at one-year follow-up. The analyses assess whether the infants differ in the proportions showing focused looks of a maximum length of > 14 seconds, or a mean length of look longer than 14 seconds. The results show no significant differences (Analysis 2.32) between the groups, indicating that the infants did not have more ability for focused attention in either group (as would be evidence by longer looks): mean looks > 14 seconds (RR 2.65; 95% CI 0.31 to 22.82); mean looks < 14 seconds (RR 0.88; 95% CI 0.68 to 1.14); maximum length of look > 14 seconds (RR 0.96; 95% CI 0.66 to 1.38); maximum length of look < 14 seconds (RR 1.76; 95% CI 0.37 to 8.31).

### Habituation

### Post-intervention

### Single study results

Cigales 1997 examined the impact of eight minutes of massage on infant habituation. Two films were repeatedly shown until infants habituated (indicated loss of interest suggesting that they had learned the colour-tempo relationships and were ready to learn something new) (Analysis 2.33; Analysis 2.34; Analysis 2.35). To examine whether the infants had habituated to these colour tempo combinations, infants then received two more trials of the same film (post-habituation trials). These indicated a non-significant difference in the time infants looked at the stimulus (MD 2.00; 95% CI -2.43 to 6.43; Analysis 2.36). Following the post-habituation trials, infants received two different test trials depicting new colour-tempo combinations and massaged infants looked significantly longer at the test trials (MD -12.40; 95% CI -19.37 to

-5.43; Analysis 2.37) compared with the post-habituation trials, suggesting that they recognised a difference in the new test trials.

## DISCUSSION

### Summary of main results

The updated review included a further 11 studies producing a total of 34 studies (Koniak-Griffin 1988, now includes data from a follow-up report) measuring the impact of infant massage on mental or physical health in typically developing infants. The number of studies and differences in outcomes necessitated that we make a number of post-hoc decisions to investigate clinical heterogeneity following meta-analyses by conducting sensitivity analyses based on risk of bias and study geographical location (East versus West). The latter was deemed to be necessary because of the diversity in terms of the usage of infant massage across these settings. We also deemed it to be worthwhile at this update to conduct subgroup analyses to investigate the effect of duration on intervention outcomes.

The 34 included studies produced a total of 14 meta-analyses for physical aspects of health measured at post-intervention (including weight, length, leg, arm, chest and head circumference, cortisol, sleep length; crying/fussing, bilirubin, incidence of illness); 18 meta-analyses of aspects of mental health (parent-infant interactions; parenting stress; attachment) and development (infant temperament, psychomotor and mental development); and three meta-analyses of weight, length and head circumference measured at follow-up. Only three meta-analyses of weight (n = 18), length (n = 11), and head circumference (n = 9) comprised five or more studies; the remaining 11 meta-analyses including data from between two and four studies. Three meta-analyses evaluated follow-up data - length, head circumference (n = 2), weight (n = 3).

Of the 14 meta-analyses assessing physical outcomes post-intervention, nine showed significant findings favouring the intervention group for weight (n = 18), length (n = 11), head circumference (n = 9), arm circumference (n = 2), leg circumference (n = 2), 24-hour sleep duration (n = 4), time spent crying/fussing (n = 4), decreased levels of blood bilirubin (n = 2), and fewer cases of diarrhoea (n = 2). Apart from one outcome (length), these significant findings were either restricted to studies at high risk of bias, or were lost following the conduct of sensitivity analyses in which studies at high risk of bias were removed.

There were no significant effects (i.e. favoured neither intervention nor control) for the following outcomes: cortisol measured at 10 to 20 minutes after a single brief intervention (n = 2), mean increase in duration of night sleep (n = 4), increase in sleep length measured over a 24-hour period (n = 2), URTI (n = 2) or anaemia (n = 2). Sensitivity analyses were conducted for weight, length and head circumference, and only the finding for length remained

significant following removal of high-risk studies. Of the three outcomes that could be meta-analysed at follow-up (i.e. length, weight and head circumference), both weight and head circumference continued to be significant at six months; however, these findings were obtained from studies conducted in Eastern countries only.

Of the 18 meta-analyses measuring aspects of mental health and development, a significant effect favouring the intervention group was found for gross motor skills (n = 2), fine motor skills (n = 2), personal and social behaviour (n = 2) and psychomotor development (n = 4). However, the first three findings were obtained from two studies, one of which was rated as being at high risk of bias, and the fourth finding was lost following a sensitivity analysis. No significant differences were found for infant temperament (three meta-analyses) (n = 3), parent-infant interaction (eight meta-analyses) (n = 2), parenting stress (n = 2), mental development (MDI) or language development (n = 2). Nor was a significant difference at follow-up for parent-infant interaction (n = 2).

Sensitivity analyses showed that all of the significant results for both physical and mental/developmental outcomes were lost once studies that were conducted in the East or that were categorised as being at high risk of bias, had been excluded, and at follow-up. The results of meta-analysis for length at post-intervention were still significant after the studies at high risk of bias due to inadequate randomisation were excluded; but all of the included studies for this analysis were carried out in the East.

The variability in the results may in part be due to the considerable heterogeneity in the studies, with an  $I^2$  close to 100% in some meta-analyses. This could in part be accounted for by differences between other study level characteristics such as setting and massage provider. However, there were no direct comparisons of types of provider or setting that would have enabled us to assess whether these factors influence the outcome. We were also unable to carry out subgroup analyses to investigate these characteristics because of variability between the studies. For example, the setting of the studies was not equivalent, in that some massage was delivered in hospitals or clinical settings; some was delivered at home or in diverse community settings; and some was delivered across a range of settings. In terms of who administered the massage (mothers or researchers/professionals), we were unable to obtain details about prior experience of massage, and how the providers were taught the massage skills, in addition to which, it was unclear whether professionals were providing massage to higher risk groups of infants compared to parents, which could further confound the analysis. We were also unable to conduct further analyses according to the massage provider because information about the identity of the provider was unclear in the published report of 12 of our included studies (carried out in China, (Wang 1999; Ke 2001; Zhai 2001; Duan 2002; Shi 2002; Sun 2004; Xua 2004; Ye 2004; Liu CL 2005; Lu 2005; Na 2005; Shao 2005), with no further details available from the trial investigators. Subgroup analyses using only those studies that provided this information, could have

introduced bias into the review methods. We were also reluctant to presuppose that identity of the provider is an important factor (it may be that the tactile stimulation is the major factor in promoting physical health measures). Finally, in Western countries, one of the primary aims of infant massage is to promote optimal parent-infant interaction (see [Background](#) for further detail), and this requires that infant massage be delivered directly by the parent.

Although we found no significant differences in terms of massage duration, the teaching sessions ranged from weekly classes of 45 to 60 minutes over four to five weeks to one demonstration and a single observation of performance. The duration and frequency of massage also varied from one episode for eight minutes to 15 minutes three times a day for six weeks. Although specific detail was often not provided, it would appear that the approach to massage also varied including the use of different massage oils in one study, tactile and kinaesthetic stimulation in another, and responsiveness to infant cues in a third. There was also considerable variation in the outcomes measures, and the measures used to assess these outcomes. These issues were reflected in the high levels of statistical heterogeneity identified in some of the meta-analyses. The conduct of post hoc subgroup analyses found no differences in outcome based on duration of intervention.

We also noted considerable variability in terms of the outcomes measures being used. For example, the impact of infant massage on sleep was assessed using duration of daytime sleep; mean increase in duration of first morning sleep; number of naps; number of hours sleep; night wake frequency; duration of night waking; 24-hour sleep duration; and increase in duration of sleep.

A number of potential biological mechanisms for an increase in growth following tactile stimulation have been identified such as for example, decrease in the growth hormone ornithine decarboxylase in rat pups removed from their mothers (Schanberg 1994); the identification of a gene underlying protein synthesis that responds to tactile stimulation (Schanberg 1994); and evidence that massage increases vagal activity which aids the secretion of gastrointestinal hormones important for food absorption, particularly insulin and gastrin, (Uvnas-Moberg 1987). However, further research is required to ascertain whether these physiological mechanisms are also evident in humans. Furthermore, the reasons for seeking an impact on outcomes such as length, weight, head/arm/leg circumference in population samples are not clear.

Similarly, evidence of significance effects of massage on catecholamine (norepinephrine and epinephrine) and cortisol excretion could potentially be very important, given what we now know about the damaging effects of high levels of stress hormones on the development of pathways in the infant brain (Gunnar 1998). Furthermore, such effects are biologically plausible - for example, tactile stimulation moderates cortisol production and promotes glucocorticoid receptors in the hippocampus (Liu 1997), although the evidence is currently limited to animals. Such an effect would also explain the potential impact of such massage on sleep and

crying. One study also reports an effect on release of melatonin (6-sulphatoxymelatonin), which is involved in the adjustment of circadian rhythms (Ferber 2002). However, the meta-analysis of outcomes for sleep was limited to a small number of studies that produced conflicting results.

There is, however, a lack of biological plausibility in terms of some of the findings. For example, Argawal 2000 suggested that the type of oil that is used is associated with the level of change identified. In this study, massaging with mustard oil improved the weight, length, and mid-arm and mid-leg circumferences as compared to infants without massage, although sesame oil was a better candidate for this than mustard oil; however, this was only one trial and the biological basis for systemic effects of different massage oils is unclear. In fact some oils such as mustard oil can have adverse effects on skin (Darmstadt 2002b).

In terms of thermal advantage, we considered if enhanced warmth resulting from massage and blood flow might contribute to improved physical outcomes, but again, the evidence for this is not available from the results of included studies and we were unable to pursue this point. We have addressed potential biological mechanisms for an increase in growth following tactile stimulation above.

Furthermore, in the absence of a significant impact on potential mediating mechanisms (for example, such as stress hormones and parent-infant interaction), it is also not clear how infant massage could impact on the many aspects of infant cognitive and developmental outcomes that are assessed in many of the included studies (for example, one study (Wang 2001) found a unusually large impact of infant massage on IQ).

### Overall completeness and applicability of evidence

Infant massage is conducted in many areas of the world and although we have endeavoured to be inclusive (we obtained evidence from both Western and Eastern countries, including India, Israel, Iran, Korea and China from the East and the UK, USA and Canada from the West), it is not clear that we have been successful in identifying all of the studies that were conducted and published in Eastern countries. Furthermore, the problems for which infant massage is delivered are wide-ranging and it is not clear that the findings of some of the included studies have universal applicability. For example, Kim 2003 found evidence of the effectiveness of infant massage in improving weight in infants living in a Korean orphanage. However, it seems possible that the biological mechanism for such an impact maybe the lack of normal stimulation that such infants receive.

### Quality of the evidence

Although the included studies were all randomised controlled trials, the quality of many was compromised by the use of quasi methods of randomisation, and many included studies also failed to specify the method of allocation concealment, and had high losses to follow-up. A large number of Eastern studies had both uniformly significant results and no reported dropout (in addition to inadequate information about their design and conduct), all of which were removed as part of sensitivity analyses. Concerns of this nature have been reported elsewhere with the recommendation to treat such data with caution (Vickers 1997). In addition, as was suggested above, despite the fact that many of the included studies examined the effect of very similar amounts and durations of massage (that is, fifteen minutes, twice daily over around six weeks), considerable statistical heterogeneity was noted, even after taking account of the individual results and the sample sizes. Selective reporting has recently been documented in other fields (for example, genetic epidemiology) of the Chinese literature, although this phenomenon may not be restricted to Chinese studies (Pan 2005). There is also documented evidence in other countries of language bias in which significant results are reported in the international literature while non-significant results appear in the local literature (Egger 1997).

### Potential biases in the review process

None known.

### Agreements and disagreements with other studies or reviews

The Vickers 2004 review of massage for promoting growth and development in pre-term or low-birth weight infants concluded that massaged babies had a weight gain of five grams a day. However, they caution against relying on this finding due to the quality of the included studies and the fact that few studies had included calorie intake. In the current review, the only evidence of any significant impact of massage on growth was similarly obtained from a group of studies regarded to be at high risk of bias. Furthermore, the use of massage to increase outcomes such as weight, length, head/arm/leg circumference is a questionable use of this intervention in population samples.

## AUTHORS' CONCLUSIONS

### Implications for practice

Infant massage is increasingly being used in the community with low-risk mother-infant dyads to promote the mother-child relationship and to improve other outcomes such as sleep. The addition of 12 new studies to this review enabled the conduct of meta-

analyses of a range of physical (for example, weight, length, head circumference, mid-thigh or leg circumference, salivary cortisol, sleep duration, mean increase in 24-hour sleep, crying or fussing time, bilirubin), mental (for example, parental stress, infant attachment, parent-infant interaction etc) and developmental (for example, temperament; physical and mental development) outcomes, of which very few achieved statistical significance, or statistical significance was lost at follow-up or following sensitivity analyses. These findings do not currently support the use of infant massage in low-risk population samples. However, the evidence that is currently available about the impact of infant massage is poor, and many studies are being conducted without addressing the biological plausibility of the outcomes being measured, the mechanisms by which change might be achieved, or indeed, the need for specific outcomes in population samples. Future research should focus on the impact of infant massage in higher-risk population samples (for example, demographically and socially deprived parent-infant dyads), where a realist evaluation has recently identified most potential for improvement (Underdown 2010).

### Implications for research

The current evidence is of poor quality and suggestive that infant massage has little impact in low-risk population samples. Further methodologically rigorous research is needed to examine the impact of infant massage on higher-risk population groups (for example, demographically and socially deprived parent-infant dyads). The evidence about the impact of compromised parent-infant interaction in terms of the infants rapidly developing neurological system is now extensive (see Background) and evaluations of appropriately focused infant massage interventions for these groups, are urgently needed.

The research should focus on the delivery of infant massage by the primary caregiver (that is, as opposed to research associates or other non-primary caregiving figures), and should be delivered routinely

for an extended period of time. So, for example, it seems likely that for infant massage to have an impact on stress hormones, it should be delivered for at least once or twice daily over a period of four to six weeks, and the integrity with which it is delivered should be monitored. Furthermore, there is evidence that for infant massage to be effective, certain mechanisms need to be present in terms of the way in which the intervention is delivered such as teaching about infant cues, optimum group size, and the setting meeting the physical needs of the client group (Underdown 2011).

There is also a need to evaluate the effectiveness of infant massage on outcomes that are biologically plausible and to identify mediatory mechanisms. For example, research evaluating the impact of infant massage on infant developmental outcomes should also measure mediatory mechanisms such as parent-infant interaction, or stress hormones (cortisol, epinephrine and norepinephrin). There is also a need for long-term follow-up to identify whether any short-term benefits that are identified are maintained over time.

### ACKNOWLEDGEMENTS

We would like to thank Yongjian Hu for his help with co-reviewing the Chinese data. Vincent Cheung identified, translated and data extracted the majority of the Chinese papers. We also thank the Cochrane CDPLPG group for their support throughout the review writing process and for the searches.

We thank the following trial investigators who responded to our requests for further information about the trial conditions: Fereshteh Narenji, Madeleine O'Higgins, Krista Oswald, Rosemary White-Traut, Xui-hong Li, Ruth Elliott and colleagues, Tiffany Field, Sari Ferber, Vonda Jump, Deborah Koniak-Griffin, Vivette Glover and Katsuno Onozawa, Tae Im Kim, Deborah Koniak-Griffin.

### REFERENCES

#### References to studies included in this review

##### Argawal 2000 *{published data only}*

Argawal K, Ashish G, Pushkarna R, Bhargava S, Prabhu M. Effects of massage and use of oil on growth, blood flow and sleep patterns in infants. *Indian Journal of Medical Research* 2000;**112**:212–7. [ : 2000]

##### Arikan 2008 *{published data only}*

Arikan D, Alp H, Gözümlü S, Orbak Z, Çifçi EK. Effectiveness of massage, sucrose solution, herbal tea or hydrolysed formula in the treatment of infantile colic. *Journal of Clinical Nursing* 2008;**17**(13):1754–61. [ : 0962–1067; : 1365–2702]

##### Cheng 2004 *{published data only}*

Cheng Z, Liu S, Zhang C, et al. Effect of touching on the development and sleeping of infants. *Journal of Nursing Science (Surgical edition)* 2004;**19**(24):10–2.

##### Cigales 1997 *{published data only}*

Cigales M, Field T, Lundy B, Cuadra A, Hart S. Massage enhances recovery from habituation in normal infants. *Infant Behavior and Development* 1997;**20**(1):29–34. [ : 1997]

##### Duan 2002 *{published data only}*

Duan Lihong, Li Weihong, Shi Fentao. Studying impact of touching on growth and development of infants. *Chinese Nursing Research* 2002;**16**:5.



- Elliott 2002** *{published data only}*  
Elliott MR, Reilly SM, Drummond J, Letourneau N. The effect of different soothing interventions on infant crying and on parent-infant interaction. *Infant Mental Health Journal* 2002;**23**(3):310–28.
- Ferber 2002** *{published data only}*  
Ferber SG, Laudon M, Kuint J, Weller A, Zisapel N. Massage therapy by mothers enhances the adjustment of circadian rhythms to the nocturnal period in full-term infants. *Developmental and Behavioral Pediatrics* 2002;**23**(6):410–5.
- Field 1996** *{published data only}*  
Field T, Grizzle N, Scafidi F, Abrams S, Richardson S. Massage therapy for infants of depressed mothers. *Infant Behaviour and Development* 1996;**19**(1):107–12.
- Jing 2007** *{published data only}*  
Jing J, Li X, Feng L, Wu Q, Wang Z, Zeng S, et al. Massage and motion training for growth and development of infants. *World Journal of Pediatrics* 2007;**3**(4):295–9.
- Jump 1998** *{published data only}*  
Jump V. *Effects of Infant Massage on Aspects of the Parent-Child Relationship: An Experimental Manipulation*. Logan, UT: Utah State University, 1998.
- Ke 2001** *{published data only}*  
Ke G, Ling X, Li L. Clinical survey of improved massage method developing infants' growth and development. *China Nurse* 2001;**36**(4):278–80.
- Kim 2003** *{published data only}*  
Kim TI, Shin YH, White-Traut AC. Multisensory intervention improves physical growth and illness rates in Korean orphaned newborn infants. *Research in Nursing & Health* 2003;**26**(6):424–33.
- Koniak-Griffin 1988** *{published data only}*  
Koniak-Griffin D, Ludington-Hoe S, Verzemnieks I. Longitudinal effects of unimodal and multimodal stimulation on development and interaction of healthy infants. *Research in Nursing and Health* 1995;**18**(1):27–38.  
\* Koniak-Griffin D, Ludington-Hoe S, Verzemnieks I. Developmental and temperament outcomes of sensory stimulation in healthy infants. *Nursing Research* 1988;**37**(2):70–6.
- Liu C 2001 0 to 2 months** *{published data only}*  
Liu Chunyang, Liu J, Lin Xiaojun. Effects of touch on growth and mentality development in normal infants. *Chinese Medical Journal (Chung-Hua i Hsueh Tsa Chih)* 2001;**81**(23):1420–3.
- Liu C 2001 3 to 6 months** *{published data only}*  
Liu Chunyang, Liu J, Lin Xiaojun. Effects of touch on growth and mentality development in normal infants. *Chinese Medical Journal (Chung-Hua i Hsueh Tsa Chih)* 2001;**81**(23):1420–3.
- Liu CL 2005** *{published data only}*  
Liu Chun Li. The effect of touching on healthy infants' weight. *Journal of Huaihai Medicine* 2005;**23**(2):137.
- Liu DY 2005** *{published data only}*  
Liu DY. Clinical experience in infant massage. 100 case studies. *Chongqing Medicine* 2005;**34**(7):301.
- Lu 2005** *{published data only}*  
Lu Jiao, Li Ju Zhan, Wu Li Fang. A study of the effect of touch on health in infants. *Practical Clinical Medicine* 2005;**6**(2):119–21.
- Maimaiti 2007** *{published data only}*  
Maimaiti R, Yan M, Muzapare D. Effect of infantile touching on behavior development in infants. *Chinese Journal of Contemporary Pediatrics* 2007;**9**(5):495–6.
- Na 2005** *{published data only}*  
Na Zhuo Hua, Xie Hui Yun, Huang Jian Hua. The effect of infant massage on growth. *New Journal of Traditional Chinese Medicine* 2005;**37**(5):69–70.
- Narenji 2008** *{published data only}*  
Narenji F, Rosbahany N. The effects of massage therapy on weight gain and sleep behaviors in infants. *Koimesh Journal of Semnan University of Medical Sciences* 2008;**9**(4):279–84. [1608–7046]
- O'Higgins 2008** *{published data only}*  
O'Higgins M. *Improving Mother-Infant Outcomes After Postnatal Depression [PhD thesis]*. London: Institute of Education, University of London, 2006.  
\* O'Higgins M, St James Roberts I, Glover V. Postnatal depression and mother and infant outcomes after infant massage. *Journal of Affective Disorders* 2008;**109**(1-2):189–92. [0165–0327]
- Onozawa 2001** *{published data only}*  
Glover V, Onozawa K, Hodgkinson A. Benefits of infant massage for mothers with postnatal depression. *Seminars in Neonatology* 2002;**7**(6):495–500.  
\* Onozawa K, Glover V, Adams D, Modi N, Kumar R C. Infant massage improves mother-infant interaction for mothers with postnatal depression. *Journal of Affective Disorders* 2001;**63**:201–7.
- Oswalt 2007** *{published data only}*  
Oswalt KL. *Outcomes of Massage Interventions on Teen Mothers and their Infants [MA thesis]*. Vol. vii, p 50, Birmingham, AL: University of Alabama at Birmingham, 2007.
- Shao 2005** *{published data only}*  
\* Shao L. The effect of infant massage on health: a nursing experience. *Journal of Chinese Clinical Research* 2005;**11**(11):1582–3.
- Shi 2002** *{published data only}*  
Shi Li, Xue Li Rong. The clinical observation of touching towards the neonatal health's affection. *Journal of Henan University* 2002;**21**(4):26–7.
- Sun 2004** *{published data only}*  
Sun Hai Yun, Gao Xiang Yu, Zhao Xue Mei. A comparative study on influence of touching on growth, development and jaundice of normal infants. *Chinese Nursing Research* 2004;**18**(8B):1468–9.

- Wang 1999** {published data only}  
Wang Bin, Shen Yue Hua, Jin Run Yan. Improving infant growth by massage: a clinical study. *Chinese Journal of Perinatal Medicine* 1999;**2**(4):241–2.
- Wang 2001** {published data only}  
Wang LX, He Q, Liu YL. The effects of infant massage on the development of infant intelligence and body weight. *Chinese Journal of Child Health Care* 2001;**6**:409–11.
- White-Traut 2009** {published data only}  
White-Traut RC, Schwertz D, McFarlin B, Kogan J. Salivary cortisol and behavioral state responses of healthy newborn infants to tactile-only and multisensory interventions. *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 2009; Vol. 38, issue 1:22–34.
- Xua 2004** {published data only}  
Xua Li Shuan, Qing Gui Romg, Ye Mei Yan, et al. Study of touch intervention of infant sleep disorder. *Journal of Nursing Continual Education* 2004;**19**(11):973–6.
- Ye 2004** {published data only}  
Ye Hong Yun. Clinical observation on infant massage. *Journal of Gan Nan Medical College* 2004;**3**:239–40.
- Zhai 2001** {published data only}  
Zhai J, Pan Xian R, Hua JR, et al. The effect of massage on growth: a clinical study. *Journal of Sichuan Continuing Education College of Public Health and Managerial Sciences* 2001;**20**(10):23.
- Zhu 2010** {published data only}  
Zhu Y. Influence of touch on neurobehavioural and intelligence development of normal infants. *China Medical Herald* 2010;**7**(21):89–91.

## References to studies excluded from this review

- Clarke 2000** {published data only}  
Clarke C, Gibb C, Hart J, Davidson A. Infant massage: developing an evidence base for health visiting practice. *Clinical Effectiveness in Nursing* 2002;**6**(3–4):121–8.
- Cullen 2000** {published data only}  
Cullen C, Field T, Escalona A, Hartshorn K. Father-infant interactions are enhanced by massage therapy. *Early Child Development and Care* 2000;**164**(1):41–7.
- Darmstadt 2002a** {published data only}  
Darmstadt G, Samir K Saha. Traditional practice of oil massage of neonates in Bangladesh. *Journal of Health, Population and Nutrition* 2002;**20**(2):184–8.
- Fernandez 1998** {published data only}  
Fernandez Y, Fernandez-Arroyo M. Study of pediatric massage [Estudio de masaje infantil]. *Anales Espanoles de Pediatria* 1998;**48**(5):549–50.
- Field 2000b** {published data only}  
Field T, Pickens J, Prodromidis M, Malphurs J, Fox N. Targeting adolescent mothers with depressive symptoms for early intervention. *Adolescence* 2000;**35**(138):381–414.
- Field 2004** {published data only}  
Field T, Hernandez-Reif M, Diego M, Feijo L, Yanexy V, Gil K. Massage therapy by parents improves early growth and development. *Infant Behavior and Development* 2004;**27**:435–42.
- Fogaça 2005** {published data only}  
Fogaça MdeC, Carvalho WB, Peres CdeA, Lora MI, Hayashi LF, Verreschi IT. Salivary cortisol as an indicator of adrenocortical function in healthy infants, using massage therapy [Fogaça Mde C, Carvalho WB, Peres Cde A, Lora MI, Hayashi LF, Verreschi IT.]. *São Paulo Medical Journal* 2005;**123**(5):215–8. [ : 1516–3180]
- Huhtala 2000** {published data only}  
Huhtala B, Lehtonen L, Heinonen R, Korvenranta H. Infant massage compared with crib vibrator in the treatment of colicky infants. *Pediatrics* 2000;**105**(6):1328.
- Im 2007** {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; Vol. 54, issue 1:31–5.
- Ineson 1995** {published data only}  
Ineson M. The psychological and physiological effects of massage on full term 'normal' babies. *Journal of the Association of Association of Chartered Physiotherapists in Women's Health* 1995;**77**:3–7.
- Jing L 2007** {published data only}  
Jing L. Effect of early environmental stimulate on infants' behaviour development. *Chinese Journal of Child Health Care* 2007;**6**:649–50.
- Jump 2006** {published data only}  
Jump VK, Fargo JD, Akers JF. Impact of massage therapy on health outcomes among orphaned infants in Ecuador: results of a randomized clinical trial. *Family & Community Health* 2006;**29**(4):314–9.
- Lee 2006** {published data only}  
Lee HK. The effects of infant massage on weight, height, and mother-infant interaction. *Journal of Korean Academy of Nursing* 2006; Vol. 36, issue 8:1331–9.
- Li 2002** {published data only}  
Xiu-hong L, Jin J, Huang J. The influence of touches on nerve and spirit and growth of normal infant. *Maternal and Child Health Care of China* 2002;**5**:299–301. [DOI: CNKI:SUN:ZFYB.0.2002-05-019]
- Oswalt 2009** {published data only}  
Oswalt KL. *Effects of Infant Massage on HIV-Infected Mothers and their Infants [PhD thesis]*. Vol. 1 online resource (vii, p 58), Birmingham, AL: University of Alabama at Birmingham, 2009.
- Pardew 1996** {published data only}  
Pardew EM. *The Effects of Infant Massage on Interactions between High Risk Infants and their Care Givers [PhD thesis]*. Corvallis, OR: Oregon State University, 1996.
- 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. *Taehan Kanho Hakhoe Chi* 2006; Vol. 36, issue 6:897–904.

**Peláez-Nogueras 1996** {published data only}

Peláez-Nogueras M, Gewirtz JL, Field T, Cigales M, Malphus J, Clasky S, et al. Infants' preference for touch stimulation in face-to-face interactions. *Journal of Applied Developmental Psychology* 1996;17(2):199–213.

**Peláez-Nogueras 1997** {published data only}

Peláez-Nogueras M, Field T, Gewirtz JL, Cigales M, Gonzalez AS, Sanchez A, et al. The effects of systematic stroking versus tickling and poking on infant attention and behaviour. *Journal of Applied Developmental Psychology* 1997;18(2):169–78.

**Peláez-Nogueras 1997b** {published data only}

Peláez-Nogueras M, Field T, Hossain Z, Pickens J. Depressed mothers' touching increases infants' positive affect and attention in still-face interactions. *Child Development* 1996;67(4):1780–92.

**Scafidi 1996** {published data only}

Scafidi F, Field T. Massage therapy improves behavior in neonates born to HIV-positive mothers. *Journal of Pediatric Psychology* 1996;21(6):889–97.

**Serrano 2010** {published data only}

Serrano MS, Doren FM, Wilson L. Teaching Chilean mothers to massage their full-term infants: effects on maternal breast-feeding and infant weight gain at age 2 and 4 months. *Journal of Perinatal & Neonatal Nursing* 2010; Vol. 24, issue 2:172–81.

**Stack 1990** {published data only}

Stack DM, Darwin W. Tactile stimulation as a component of social interchange: new interpretations for the still-face effect. *British Journal of Developmental Psychology* 1990 June;8(2):131–45.

**Yilmaz 2009** {published data only}

Yilmaz HB, Conk Z. The effect of massage by mothers on growth in healthy full term infants. *International Journal of Human Sciences* 2009;6(1):970–7.

**Zhu 2000** {published data only}

Zhu JX, Wu SM, Wang DF, Yang HW, Shen YH, Wu W, et al. The effects of infant touch on neonatal growth and development: a multicenter clinical trial. *Journal of Applied Clinical Pediatrics* 2000;15:4.

**Zhu 2002** {published data only}

Zhu S, Sun Ji, Ma Y, Zhang G, Yu L, Lin Y. Experimental research on promoting growth and child health care by massage and caress. *Chinese Journal of Child Health Care* 2002;1:38–40.

## Additional references

**Arnott 2007**

Arnott B, Meins E. Links between antenatal attachment representations, postnatal mind-mindedness, and infant attachment security: a preliminary study of mothers and fathers. *Bulletin of the Menninger Clinic* 2007;71:132–49.

**Bayley 1969**

Bayley N. *Bayley Scales of Infant Development*. New York: Psychological Corporation, 1969.

**Beebe 2010**

Beebe B, Jaffe J, Markese S, Buck K, Chen H, Cohen P, et al. The origins of 12-month attachment: a microanalysis of 4-month mother-infant interaction. *Attachment and Human Development* 2010;12(1-2):3–141.

**Bradley 1977**

Bradley RH, Caldwell B. Home observation for measurement of the environment: a validation study of screening efficiency. *American Journal of Mental Deficiency* 1977;81:417–20.

**Bredy 2003**

Bredy TW, Grant RJ, Champagne DL, Meaney MJ. Maternal care influences neuronal survival in the hippocampus of the rat. *European Journal of Neuroscience* 2003;18(10):2903–9.

**Darmstadt 2002b**

Darmstadt GL, Mao-Qiang M, Chi E, Saha SK, Ziboh VA, Black RE, et al. Impact of topical oils on the skin barrier: possible implications for neonatal health in developing countries. *Acta Paediatrica* 2002;91(5):546–54.

**De Wolff 1997**

De Wolff MS, Van Ijzendoorn MH. Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment security. *Child Development* 1997;68(4):604–9.

**Degangi 2000**

Degangi GA, Breinbauer C, Doussard-Roosevelt J, Porges S, Greenspan S. Prediction of childhood problems at three years in children experiencing disorders of regulation during infancy. *Infant Mental Health Journal* 2000;21(3):156–75.

**Degnan 2008**

Degnan K, Calkins S, Keane S. Profiles of disruptive behavior across early childhood: contributions of frustration reactivity, physiological regulation, and maternal behavior. *Child Development* 2008;79(5):1357–76.

**Egger 1997**

Egger M. Language bias in randomised controlled trials published in English and German. *Lancet* 1997;350(9074):326.

**Field 1996b**

Field T, Schanberg S, Davalos M, Malphurs J. Massage with oil has more positive effects on normal infants. *Journal of Prenatal and Perinatal Psychology and Health* 1996;11(2):75–80.

**Fonagy 2002**

Fonagy P, Gergely G, Jurist E, Target M. *Affect Regulation, Mentalization and the Development of the Self*. London: Karnac, 2002.

**Fonagy 2007**

Fonagy P, Gergely G, Target M. The parent-infant dyad and the construction of the subjective self. *Journal of Child Psychology and Psychiatry* 2007;48:288–328.

**Gianino 1988**

Gianino A, Tronick E. The Mutual Regulations Model: the infant's self and interactive regulation, coping and defense.

- In: T Field, P McCabe, N Schneiderman editor(s). *Stress and Coping*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988:47–68.
- Gunnar 1998**  
Gunnar M. Quality of early care and buffering of neuroendocrine stress reactions: potential effects on the developing human brain. *Preventive Medicine* 1998;**27**: 208–11.
- Gunnar 2007**  
Gunnar M, Quevedo K. The neurobiology of stress and development. *Annual Review of Psychology* 2007;**58**:145–73.
- Higgins 2002**  
Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**:1539–58.
- Higgins 2011**  
Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org) 2011.
- Ireland 2000**  
Ireland M, Olson M. Massage therapy and therapeutic touch in children: state of the science. *Alternative Therapies in Health and Medicine* 2000;**6**(5):54–60.
- Johnson 2011**  
Johnson, Johnson Ltd. Guide to Infant Massage. <http://www.johnsonsbaby.co.uk/massage/guides/newborn-guide-to-infant-massage/2> (accessed 5 October 2011).
- Keren 2001**  
Keren M, Feldman R, Tyano S. Diagnoses and interactive patterns of infants referred to a community based infant mental health clinic. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001;**40**:27–35.
- Kersten-Alvarez 2011**  
Kersten-Alvarez LE, Hosman CMH, Riksen-Walraven JM, Van Doesum KTM, Hoefnagels C. Which preventive interventions effectively enhance depressed mothers' sensitivity? A Meta-analysis. *Infant Mental Health Journal* 2011;**32**(3):362–76.
- Koniak-Griffin 1995**  
Koniak-Griffin D, Ludington-Hoe S, Verzemnick I. Longitudinal effects of unimodal and multimodal stimulation on development and interaction of healthy infants. *Research in Nursing and Health* 1995;**18**(1):27–38.
- Kropp 1987**  
Kropp J, Haynes O. Abusive and non-abusive mothers' ability to identify general and specific emotion signals of infants. *Child Development* 1987;**58**(1):187–90.
- Liu 1997**  
Liu D, Diorio J, Tannenbaum B. Maternal Care, hippocampal glucocorticoid receptors and hypothalamic-pituitary-adrenal responses to stress. *Science* 1997;**277**: 1659–62.
- Liu 2001**  
Liu C, Liu J, Lin X. Effects of touch on growth and mentality development in normal infants. *Chinese Medical Journal (Chung-Hua i Hsueh Tsa Chih)* 2001;**81**(23): 1420–3.
- Liu 2002**  
Liu D, Diorio J, Day JC, Francis DD, Meaney MJ. Maternal care, hippocampal synaptogenesis and cognitive development in rats. *Nature and Neuroscience* 2000;**3**(8): 799–806.
- Medoff-Cooper 1993**  
Medoff-Cooper B, Carey W, McDevitt S. The early infancy temperament questionnaire. *Journal of Developmental and Behavioral Pediatrics* 1993;**14**(4):230–5.
- Murray 1992**  
Murray L. The impact of post natal depression on infant development. *Journal of Child Psychology and Psychiatry* 1992;**33**(3):543–61.
- Murray 1996**  
Murray L, Fiori-Cowley A, Hooper R, Cooper P. The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. *Child Development* 1996;**67**:2512–26.
- Pan 2005**  
Pan Z, Trikalinos TA, Kavvoura FK, Lau J, Ioannidis J. Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese Literature. *PLoS Medicine* 2005;**2** (12):1309–17.
- Papousek 1995**  
Papousek M, Hofacker N. Persistent crying and parenting: search for a butterfly in a dynamic system. *Early Development and Parenting* 1995;**4**(4):209–24.
- RevMan 5.1.7**  
Review Manager. RevMan. 5.1.7. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011, 2011.
- Robinson 1980**  
Robinson EA, Eyberg SM, Ross AW. The standardization of an inventory of child conduct problem behaviours. *Journal of Clinical and Child Psychology* 1980;**9**:22–8.
- Rowe 1977**  
Rowe DC, Plomin R. Temperament in early childhood. *Journal of Personality Assessment* 1977;**41**:150–6.
- Schanberg 1994**  
Schanberg S. Genetic basis for touch effects. In: Field, T editor(s). *Touch in Early Development*. Hillsdale, New Jersey: Erlbaum, 1994:67–80.
- Shai 2011**  
Shai D, Belskey J. When words just won't do: introducing parental embodied mentalising. *Child Development Perspectives* 2011;**5**(3):173–80.
- Thoman 1981**  
Thoman EB, Deenberg VH, Sieval J, Zeidner L, Becker PT. State organisation in neonates: developmental inconsistency indicates risk for developmental dysfunction. *Neuropaediatria* 1981;**12**(1):45–54.

**Thoman 1987**

Thoman EB. The sleeping and waking states of infants: correlations across time and person. *Physiology and Behaviour* 1987;**41**:531–7.

**Tronick 1982**

Tronick E, Ricks M, Cohn J. Maternal and infant affective exchange: patterns of adaptation, emotion and early interaction. In: Field T, Fogel A editor(s). *Emotion and Early Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1982.

**Tronick 1989**

Tronick EZ. Emotions and emotional communication in infants. *American Psychologist* February 1989;**44**(2):112–9.

**Tronick 2007**

Tronick E. *The Neurobehavioral and Social-Emotional Development of Infants and Children*. New York: Norton and Co, 2007.

**Underdown 2010**

Underdown A. A Realist Evaluation of the Processes and Outcomes of Infant Massage Programmes. Doctoral Thesis: University of Warwick 2010.

**Underdown 2011**

Underdown A, Barlow J. Interventions to support early relationships: mechanisms identified within infant massage programmes. *Community Practitioner* 2011;**84**(4):21–6.

**Uvnas-Moberg 1987**

Uvnas-Moberg K, Widstrom AM, Marchini G, Windberg J. Release of GI hormones in mother and infant by sensory

stimulation. *Acta Paediatrica Scandinavica* 1987;**76**(6): 851–60.

**Vickers 1997**

Vickers A, Niraj G, Harland R, Rees R. Do certain countries produce only positive results? A systematic review of controlled trials. *Controlled Clinical Trials* 1998;**19**:159–66.

**Vickers 2004**

Vickers A, Ohlsson A, Lacy JB, Horsley A. Massage for promoting growth and development of preterm and/or low birth-weight infants. *Cochrane Database of Systematic Reviews* 2004, Issue 4. [DOI: 10.1002/14651858.CD000390.pub2]

**Waters 1985**

Waters E, Deane KE. Defining and assessing individual differences in attachment relationships: Q-methodology and the organization of behavior in infancy and early childhood. *Monographs of the Society for Research in Child Development* 1985;**50**(1-2):41–65.

**Zeanah 2000**

Zeanah CHJ (editor). *Handbook of Infant Mental Health*. 2nd Edition. New York: Guilford Press, 2000.

**References to other published versions of this review****Underdown 2006**

Underdown A, Barlow J, Chung V, Stewart-Brown S. Massage intervention for promoting mental and physical health in infants aged under six months. *Cochrane Database of Systematic Reviews* 2006, Issue 4. [DOI: 10.1002/14651858.CD005038.pub2; : CD005038]

\* Indicates the major publication for the study

## CHARACTERISTICS OF STUDIES

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

#### Argawal 2000

Methods	Design: randomised controlled trial. Setting: community clinic, India.
Participants	125 healthy infants, n = 25 in each group, 6 weeks +/- 1 week of age
Interventions	Massage infants received (i) herbal oil, (ii) sesame oil, (iii) mustard oil, or (iv) mineral oil for massage daily over four weeks versus a 'no treatment' control group Massage provider: mothers trained by researchers. Duration of intervention: daily for 4 weeks for 10 minutes each session (short duration of intervention)
Outcomes	Anthropometric measurements: microhaematocrit; serum proteins, creatinine and creatine phosphokinase; blood flow using colour doppler Sleep pattern; weight (kg); length (cm); head Circumference (cm); Mid-arm circumference (cm); Mid-leg circumference (cm); Microhaematocrit; Serum proteins; Serum albumin; Serum creatinine; Creatinine phosphokinase.
Notes	Funder not stated.

#### *Risk of bias*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table divided into 5 groups, n = 25 in each group
Allocation concealment (selection bias)	High risk	Inadequate.
Incomplete outcome data (attrition bias) All outcomes	Low risk	None dropped out. Attendance strictly regulated with mothers attending weekly to have their massage techniques monitored and to return empty oil bottles before collecting their next week's supply of specific oils
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.

**Argawal 2000** (Continued)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Limited to supplying oil in different opaque bottles on different days. Key to oils was opened only at the end of the study

**Arikan 2008**

Methods	Design: randomised controlled trial. Setting: community. Public healthcare clinics and Dept of Pediatrics of Yakutiye Research Hospital, Turkey
Participants	Sample sizes: 175 infants with diagnosed colic (Wessel) randomised into 35 per group of massage, versus sucrose solution, versus herbal tea versus hydrolysed formula versus control Ages: intervention 2.29 months SD 0.75; Control 2.28 months SD 0.61 Gender: intervention 46% boys; control 34% boys. Massage provider: mothers trained by researchers.
Interventions	Mothers were trained in massage technique and given brochures with written illustrated instructions Massage (“chiropractic spinal manipulation”), twice a day for 25 minutes duration during symptoms of colic for one week (short duration of intervention) Control group: no treatment. Massage provider: mothers.
Outcomes	Parent report using daily structured diary, onset of crying time, when the intervention was given, cessation of crying time, any side effects Crying was quantified by length of crying in hours per day for one week before and one week during the intervention Timing: outcomes assessed after one week of intervention.
Notes	Funder: “we did not receive any financial support for this study” p. 1760

***Risk of bias***

<b>Bias</b>	<b>Authors’ judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	Unclear risk	Described as randomised and controlled, but no details given. No further information available from investigator
Allocation concealment (selection bias)	Unclear risk	Unclear, no details given.

**Arikan 2008** (Continued)

Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropouts and losses to follow-up not stated but intervention was only one week. Results for 35/35 reported for each of the 4 groups (Table 3 p. 1759).
Selective reporting (reporting bias)	Low risk	Crying time as the only outcome reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote “because of the design of the study, blinding was not possible”
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: blinding not possible and the same paediatrician and nurse were in contact with all study parents Comment: parent self report (diary) of crying time.

**Cheng 2004**

Methods	Design: randomised controlled trial. Setting: primary care (post-natally in hospital then in community)
Participants	Sample sizes: n = 100; intervention n = 50; control n = 50. Ages: one day after birth then daily until 42 days. Gender: in total sample 54% male, 46% female.
Interventions	Duration, dose, type. 15 min once daily for 42 days versus routine (no massage) care (medium-term duration of intervention) Massage provider: mothers.
Outcomes	Types of outcome: head circumference, length, weight, sleep duration and crying time Timing of assessment: at 3 days and 42 days of age.
Notes	Funder: not stated.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Described as ‘randomly divided’ but no details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given.



**Cheng 2004** (Continued)

		Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	100/100 results reported. No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Methods (article keywords) describe a 'blind' study but it is unclear who was blinded and how

**Cigales 1997**

Methods	Design: randomised controlled trial. Setting: hospital (research clinic), USA.
Participants	56 4-month-old infants recruited, n = 20 massage, n = 12 no stimulation control group
Interventions	Massaged infants were given either as single session of 8 minutes of massage, play, versus a no stimulation control group prior to an audiovisual habituation task (brief duration of intervention) Massage provider: investigator.
Outcomes	Average number of seconds of looking on two post habituation trials (PH) and two test trials (T) to yield a post habituation score and a test score
Notes	Funder: Johnson and Johnson.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	Unclear risk	No details given.
Allocation concealment (selection bias)	Unclear risk	Unclear.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unclear. Unclear which groups dropouts came from ( p. 30 "20 further infants were excluded from the study due to excessive crying or fussing n = 12, falling asleep n = 3; experimenter error n = 4 and fatigue n

**Cigales 1997** (Continued)

		= 1"). Results reported for n = 20 in the massage group, n = 24 in the control 'play' group, n = 12 no stimulation control group
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote p. 31 "a second observer who was blind to the pre-habituating treatment of the infants coded the visual fixations of 40% of the sample from the pre-recorded videos". Comment: unclear if complete blinding was achieved

**Duan 2002**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	160 newborn infants (n = 80 massage, n = 80 control).
Interventions	Massaged for 15 minutes twice daily over 42 days versus a 'routine care' control group (medium-term duration of intervention) Massage provider: unclear.
Outcomes	Weight, length and head circumference.
Notes	Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Unclear. Described as randomised but no details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given. Comment: judged as high risk, no further details available from trial investigator

**Duan 2002** (Continued)

Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	Unclear, outcomes of interest not prespecified in this short report; no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded at all.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not blinded at all.

**Elliott 2002**

Methods	Design: randomised controlled trial. Setting: community after training of parents to carry out massage, Canada
Participants	111 first time parent-infant dyads (newborns).
Interventions	<p>Group 1: Massage (n = 31)            Group 2: Supplemental carrying (n = 29)            Group 3: Massage and supplemental carrying (n = 24)            Group 4: No treatment control group (n = 27)</p> <p>Massage group:            7 - 10 days postpartum parents taught massage plus they received a video tape showing the steps and printed instructions. Minimum of 10 mins daily, up to 20 mins daily, 2 to 16 weeks of age (long duration of intervention).            2nd home visit parent was assessed by research assistant (RA) to check that massage covered 85% of infant's body and took 10-20 mins. to complete.</p> <p>Supplemental carrying group:            Received carrier and instructions for use.            Carried infant in carrier for minimum of 3 hours not only in response to crying but in addition to time spent feeding and independent of whether the infant was awake or asleep.</p> <p>Supplemental carrying/massage:            Received instruction and equipment for both interventions above</p> <p>Massage provider: mothers trained by researchers.</p>
Outcomes	<ol style="list-style-type: none"> <li>1.Nursing Child Assessment Sleep Activity Record (NCASA)</li> <li>2.Nursing Child Assessment Feeding (NCAFS) and teaching</li> <li>3.Early infant temperament questionnaire (EITQ)</li> <li>4.State Trait Anxiety Inventory - STAI-T-anxiety scale</li> <li>5.Parental sense of competence scale (PSOC)</li> <li>6.Difficult life circumstances scale (DLC)</li> </ol>

Notes	Funders: Canadian Nurses' Foundation, the Alberta Foundation for Nursing Research, the University of Alberta, and the University of Calgary	
<b>Risk of bias</b>		
<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	Low risk	Described as randomised, used repeated measures design involving a randomised two-way layout with treatment factors 'carrying' and 'massage' as two levels to ensure that every dyad had an equal chance of being assigned to one of four groups
Allocation concealment (selection bias)	Low risk	p. 319 "Research associate, who was not involved with the subjects randomly assigned subjects to one of the four groups whenever a subject agreed to enter the study."
Incomplete outcome data (attrition bias) All outcomes	Low risk	94/111 were present at week 16. Reasons for dropout or loss to follow-up given. From p. 321 "One infant was intolerant of the intervention (massage), 5 withdrew because they no longer met criteria as infants required hospital, 1 infant still born 4 left family issues 7 study too time consuming Dropouts occurred across all four study groups: Massage (n = 6), Supplemental carrying (n = 3), Supplemental carrying/massage (n = 3), Control (n = 5)."
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Parents and outcome data collectors were kept apart to insure that observed differences occurred as a result of the treatment

**Ferber 2002**

Methods	Design: randomised controlled trial. Setting: community after training of parents to carry out massage, Israel
Participants	21 dyads of mothers and full term infants (n = 13 massage; n = 8 control)
Interventions	Massage provider: mothers trained by researchers. Massage therapy was performed daily by the mother for 14 days versus a no treatment control group (short duration of intervention)
Outcomes	1. Circadian rhythmicity 2. Excretion of the main melatonin metabolite 6sulphatoxymelatonin
Notes	Funder: The Academic Research Funds and the Social Science Dean's scholarships at Bar-Ilan University

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	Unclear risk	Unclear.
Allocation concealment (selection bias)	Unclear risk	Unclear.
Incomplete outcome data (attrition bias) All outcomes	Low risk	52 mothers were asked to participate with their babies within 2-3 days post partum. Of this group 50% (n = 26) agreed, 19.9% (n = 5) discontinued after first measurements Reported a dropout rate of 20% with no significant differences between the two intervention and control groups;
Selective reporting (reporting bias)	Low risk	Measurements for sleep rest activity from the actigraph were quoted as mean movement scores using a graph but do not have SDs. Means and SDs later supplied by investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Unclear. Actigraph measurements and the 6-sulphatoxymelatonin secretions were analysed separately but does not clarify whether the assessors were blind to the participant group

## Field 1996

Methods	Design: quasi-randomised controlled trial. Setting: community (daycare - nursery school), USA.	
Participants	40 full-term 1 - 3 month old infants, recruited if their adolescent mothers were diagnosed as depressed following delivery. n = 20 massage; n = 20 control	
Interventions	Infants in the intervention group received massage by a researcher (complete face and body using mineral baby oil); the control group infants were rocked (by cradling in the arms of the researcher). Massage delivered for 15 mins a day 2 days a week over 6 weeks Massage provider: researchers.	
Outcomes	1.Sleep/wake behaviours (Thoman 1981). 2.Salivary cortisol (ng/mL). 3.Weight (lb) and formula intake (volume, no units given, assumed US fl. oz). 4.Temperament ratings - using Colorado Child Temperament Scale	
Notes	NIMH grants and Johnson and Johnson, Gerber Foundation.	
<b><i>Risk of bias</i></b>		
<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Quasi-randomised, but no additional details of how infants were selected were provided
Allocation concealment (selection bias)	Unclear risk	Unclear, no details given.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropout for 40 post-natally depressed mother-infant dyads because the infants were being cared for by teachers in a nursery school during the six-week study (medium-term duration of intervention)
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention. Researchers carried out the massage
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Teachers and mother who recorded outcomes were unaware of therapy (mothers) or intent of study (teachers). Comment some attempt at blinding was made although it is implied that teacher knew which therapy was delivered

**Jing 2007**

Methods	Design: randomised controlled trial. Setting: community. Research clinic affiliated to Sun Yat-sen University, China
Participants	Sample sizes: n = 180 intervention n = 90; control n = 90. Ages: from birth 0 months group, the 6 month group was excluded as they are outside our age inclusion criteria Gender: not stated.
Interventions	Motion training, including gross motion and fine motion, was performed on the basis of Johnson infant massage A set of training programmes adapted to the age and development of infants was used (no details given) In the experimental group, the parents of the infants were trained to massage and motion training. All the parents were given manuals and VCD to learn the procedures Massage and motion training was performed 1-2 times every day, lasting for 15 minutes, and motion training for 5 minutes at each time, from birth to 6 months of age. From 6 months of age massage and motion training continued (massage 8 mins, motion training 12 mins). Motion training is included in the Johnson massage method (long duration of intervention) Control group received no intervention (details from trial investigator) Massage provider: trained parents.
Outcomes	Weight (kg) at 0, 1, 6 and 12 months. Length (cm) at 0, 1, 6 and 12 months. (Also weight and length enhancement using comparisons from baseline to 6 months and 6 to 12 months) Developmental quotient (Gesell Developmental Schedule) at 1, 6 and 12 months
Notes	Funder: declared as 'none'.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	Low risk	Described as randomised p. 286 (by random numbers table, further information from trial investigator)
Allocation concealment (selection bias)	High risk	Not concealed (further information from trial investigator).
Incomplete outcome data (attrition bias) All outcomes	High risk	From birth group: 54/90 at one year in the intervention group, 62/90 in the control group It is unclear how many infants were lost to follow-up at the 6 month (post-intervention time point) Losses to follow-up are reported at post-intervention, but the reasons for loss to fol-

**Jing 2007** (Continued)

		low-up are unclear (no details given, trial investigators contacted but don't know the reasons)
Selective reporting (reporting bias)	Low risk	Length, body weight and developmental quotient (Gesell DQ) pre-specified as outcomes and reported
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention. Mother in the intervention group knew their allocation, the mothers in the control group did not know the allocation (further information from trial investigator)
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Unclear if outcome assessors were blinded.

**Jump 1998**

Methods	Design: quasi-randomised controlled trial. Setting: community after training of parents to carry out massage (parenting class), USA	
Participants	57 mother-infant dyads with babies under 9 months intervention n = 27; control n = 30	
Interventions	Mothers trained in the use of infant massage in groups delivered over 45 to 60 minute sessions once a week over 4 weeks (participants encouraged to practice massage on their infants daily in between sessions), plus information about infant development (short duration of intervention) Control group received information about infant development only	
Outcomes	<ol style="list-style-type: none"> <li>1. Attachment Q set scored as a continuous variable</li> <li>2. Parenting stress Index PSI (Abidin 1986) - child and parent variables were analysed separately in their respective scales as well as combined into composite child and parent scores</li> <li>3. Adult attachment style was measured using the relationship survey</li> <li>4. Infant temperament was measured using the Infant Behaviour Questionnaire (IBQ)</li> <li>5. Parental attitudes were measured using the parental attitudes toward child rearing (PACR)</li> </ol> Massage provider: mothers trained by researchers.	
Notes	Funder: unclear.	

***Risk of bias***

Bias	Authors' judgement	Support for judgement
------	--------------------	-----------------------



**Jump 1998** (Continued)

Random sequence generation (selection bias)	High risk	Quasi-randomised. Used coin flip to assign the first infant and the remaining infants were alternatively allocated to the intervention or control group
Allocation concealment (selection bias)	Unclear risk	Unclear.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	12/57 lost over 12 months with no forwarding information. 21 in final intervention group 24 in final control group 21% dropout rate; mothers from both groups who left the study were less educated and had younger infants than those remaining in the study, the groups were otherwise alike demographically
Selective reporting (reporting bias)	Unclear risk	A 'battery' of questionnaires was given to all participants before the intervention plus demographic details - after the 4-week period another 'battery' of questionnaire were given to all participants but these data are not part of this study page 45 (only data at 12 months is reported)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not done due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding of outcome assessors.

**Ke 2001**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	400 newborn infants intervention n = 200; control n = 200.
Interventions	Fifteen minutes of massage three times a day for 42 days plus additional method of kneading the back versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: unclear.
Outcomes	Weight, length and head circumference. Additional measures included grasp of hands, stretch and crook of front arms etc

**Ke 2001** (Continued)

Notes	Funder: unclear	
<b>Risk of bias</b>		
<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given. Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	Unclear. No outcomes pre-specified. Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Kim 2003**

Methods	Design: quasi-randomised controlled trial. Setting: orphanage, Korea.
Participants	58 Korean orphaned infants, within 14 days of birth. Intervention n = 30, control n = 28
Interventions	In addition to receiving the routine orphanage care, infants in the experimental group received 15 min twice a day of auditory (female voice), tactile (massage), and visual (eye-to-eye contact) stimulation for 4 weeks, versus a 'usual orphanage care' control group (short duration of intervention) Massage provider: researchers/orphanage staff.
Outcomes	Weight Head circumference Length

**Kim 2003** (Continued)

Notes	Also presents results for six-month follow-up. Funder: unclear.	
<b>Risk of bias</b>		
<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	First infant assigned by flip of coin then alternately after that. Quasi-randomised
Allocation concealment (selection bias)	Unclear risk	Unclear.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	At 6 months 13/58 infants had been lost to the trial because they were adopted (22%) . The loss was evenly spread between the groups, impacting on the power of the study
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Unclear, the outcome of "illness" was assessed by a nurse who was blind to the infant group assignments. It is unclear if the other outcomes were assessed blindly Quote p. 431 "although precautions were taken to keep the orphanage staff blind to group assignment (staff members were out of the room during the intervention period), the staff (including the nurse who assessed infant illness) may have been aware of group assignment"

**Koniak-Griffin 1988**

Methods	Design: randomised controlled trial and 24-month follow-up. Setting: two community hospitals in Southern California, USA
Participants	81 primiparous mothers and newborn infants (3rd or 4th days after birth), data for 49 of original 81 infants reported at follow-up
Interventions	1. The unimodal stimulation group received infant massage (5-7 minutes) once daily (n = 20). 2. The multimodal stimulation group of infants were placed on a hammock with multi-sensory elements during expected sleep periods. A simulated heartbeat and mild vestibular

**Koniak-Griffin 1988** (Continued)

lar stimulus were added continuously during the sleep period (n = 20).  
 3. The combined stimulation group received both (n = 20).  
 4. No treatment control group (n = 21).  
 All interventions were given until the infants reached 3 months of life (medium-term duration of intervention)  
 Massage provider: mothers trained by researchers.

**Outcomes**  
 Weight (g)  
 Bayley Scales of Infant Development (BSID)  
 Eyberg's Child Behavior Inventory  
 Nursing Child Assessment Teaching Scales (NCATS)  
 Revised Infant Temperament Questionnaire (RITQ) \* high score worse  
 At 24 months follow-up:  
 1. Bayley Scales of Infant Development (BSID)  
 2. Eyberg's Child Behavior Inventory  
 3. Nursing Child Assessment Teaching Scales (NCATS)  
 4. HOME Inventory

**Notes**  
 Also presents results for eight-month follow-up.  
 Funder: University of California.  
 Follow-up funder: University of California and Cuddle International

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	Unclear risk	Unclear: states 'randomly assigned'.
Allocation concealment (selection bias)	Unclear risk	Unclear.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Data for only 41 children at 4, 8 and 24 months representing an attrition rate of 39%, due to families moving out of the area. In the follow-up study, data were shown at 4 and 8 months only for those 41 infants who had completed the study at 24 months (further information from study investigator)
Selective reporting (reporting bias)	Unclear risk	Although all 3 components of the Bayley scales of infant development were administered at 4 and 8 months of age, only findings related to MDI and PDI are presented
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.

**Koniak-Griffin 1988** (Continued)

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Independent assessors (nurses) who were blind to group assignment (page 72)
---	----------	---

**Liu C 2001 0 to 2 months**

Methods	Design: randomised controlled trial. Setting: China, community (specialised massage clinic and at home)	
Participants	Sample sizes: n = 232; intervention n = 159; control n = 73. Ages: not stated, 0-2 months. Gender: not stated.	
Interventions	Massage 2 -3 times daily for 15 mins for at least 3 months (medium-term duration of intervention). Massage method by Johnson and Johnson. Carried out by parents who were first trained by doctors at a specialist massage centre. Telephone support and contact from doctors in first month As touch group but without massage - treatment as usual. Massage provider: mothers trained by researchers.	
Outcomes	Primary outcome data (Means and SDs): Bayley MDI mental development index, Bayley PDI psychological development index, sleep habits (good, not good, medium), growth (height, weight head circumference, chest circumference) statistical significance only using T and p. Illness (URTI, diarrhoea, anaemia) Timing: outcomes assessed at baseline and at 6 months from start of intervention	
Notes	Funder: not stated. If the babies developed anaemia during the studies they were treated with oral iron supplementations until the Hb levels reached normal and then for one month after	

***Risk of bias***

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Described in abstract of study as "randomly divided", but no details given Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report

**Liu C 2001 0 to 2 months** (Continued)

Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded, not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding of outcome assessors.

**Liu C 2001 3 to 6 months**

Methods	Design: randomised controlled trial. Setting: China, community (specialised massage clinic and at home)	
Participants	Sample sizes: n = 78; intervention n = 49; control n = 29 Ages: not stated 3-6 months Gender: not stated	
Interventions	Massage 2 -3 times daily for 15 min for at least 3 months (medium-term duration of intervention). Massage method by Johnson and Johnson. Carried out by parents who were first trained by doctors at a specialist massage centre. Telephone support and contact from doctors in first month As touch group but without massage - treatment as usual. Massage provider: parents	
Outcomes	Primary outcome data (Means and SDs): MDI mental development index, PDI psychological development index, sleep habits (good, not good, medium), growth (height, weight head circumference, chest circumference), statistical significance only using T and p. Illness (URTI, diarrhoea, anaemia) Timing: outcomes assessed at baseline and at 6 months from start of intervention	
Notes	Funder: not stated If the babies developed anaemia during the studies they were treated with oral iron supplementations until the Hb levels reached normal and then for one month after	

***Risk of bias***

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Described in abstract of study as "randomly divided", but no details given Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further

**Liu C 2001 3 to 6 months** (Continued)

		details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded, not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding of outcome assessors.

**Liu CL 2005**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	80 newborn infants: n = 40 intervention; n = 40 control.
Interventions	15 minutes of massage twice daily over 42 days versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: unclear.
Outcomes	Weight.
Notes	Paper not fully transcribed. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report

**Liu CL 2005** (Continued)

Selective reporting (reporting bias)	High risk	Outcomes not clearly pre-specified, weight cannot be assessed as there is no description of the measurement methods Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Liu DY 2005**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	200 newborn infants: n = 100 intervention; n = 100 control.
Interventions	15 minutes of massage twice daily over 42 days carried out by nurses versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: nurses.
Outcomes	Weight, height, head circumference and length of sleep.
Notes	Paper not fully transcribed. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report



**Liu DY 2005** (Continued)

Selective reporting (reporting bias)	High risk	Sleep length cannot be assessed as there is no description of the measurement methods Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Lu 2005**

Methods	Design: quasi-randomised controlled trial. Setting: unclear, China.
Participants	200 newborn infants: n = 100 intervention; n = 100 control.
Interventions	15 minutes of massage twice daily over 3 months versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: unclear.
Outcomes	Weight, height, head circumference and bilirubin.
Notes	Paper not fully transcribed.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Quasi-randomised according to sequence of birth dates.
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given. Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	High risk. Neural function and development was assessed but the outcome mea-

**Lu 2005** (Continued)

		sure was not validated Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Maimaiti 2007**

Methods	Design: randomised controlled trial. Setting: community, China.
Participants	Sample sizes: n = 200; intervention n = 100; control n = 100 Ages: intervention began one day after birth. Ages not otherwise stated Gender: intervention 55% male; control 53% male.
Interventions	3 times daily by trained professional while in hospital starting one day after birth then parents trained to continue massage once discharged from hospital. Duration not stated No intervention for control group. Massage provider: professionals initially then trained parents
Outcomes	Infant physical development characteristics including angle at which the infant can rise from a prone position, sight and auditory tracking and ability to smile. Statistical significance only of weight, length and head circumference are reported using X <sup>2</sup> -sided test with P values. Timing: unclear.
Notes	Funder: not stated.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator

**Maimaiti 2007** (Continued)

Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	Weight, length and head circumference were also measured but only reported as being significant
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Na 2005**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	80 newborn infants; n = 40 intervention; n = 40 control.
Interventions	15 minutes of massage three times daily for 28 days versus a 'no treatment' control group (short duration of intervention) Massage provider: unclear.
Outcomes	Weight, height, head circumference.
Notes	Paper not fully transcribed. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study re-

Na 2005 (Continued)

		port
Selective reporting (reporting bias)	High risk	No description of measurement methods, for physical growth results Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Narenji 2008**

Methods	Design: randomised controlled trial. Setting: community (clinic based), Iran.
Participants	Sample sizes: n = 100; intervention n = 50; control n = 50. Ages: infants aged 2 months no further details given. Gender: no statistical differences in gender (or other characteristics) at start of study, but no further details given
Interventions	Mothers trained to massage babies, massage all over the body excluding the eyes and genitals using sesame oil. Twice daily for 10 mins, for 4 weeks (morning and night before sleep) (short duration of intervention) Massage provider: mothers trained by researchers.
Outcomes	Weight, height, head circumference, chest circumference, abdominal circumference, arm length, thigh circumference Sleep duration in 24 hours before study and at outcome assessment Number of hours slept at night before and after study. Timing: at 4 weeks.
Notes	Funder: supported by the University of Arak (training and research assistance)

***Risk of bias***

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Described as randomly assigned (by random numbers table (further information from trial investigator) Infants were randomly assigned to one of two clinics.

**Narenji 2008** (Continued)

Allocation concealment (selection bias)	Low risk	By sealed envelope, (further information from trial investigator)
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropout or loss to follow-up. Number of participants in each group n = 50 (further information from trial investigator)
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The outcome assessors did not know whether the infants received massage or no massage as the infants were identified by a coded number (further information from trial investigator)

**O'Higgins 2008**

Methods	Design: quasi-randomised controlled trial (randomised on the basis of timing of intervention) Setting: community classes, UK.
Participants	Sample sizes: n = 96; intervention n = 45; control n = 51. Ages: intervention 9 weeks of age (median); control 10 weeks of age (median) Gender: intervention 45.2% male; control 48.4% male. (Mothers who provided massage were recruited from a group with depressive symptoms)
Interventions	Duration 1h. Frequency ideally one session per week if possible. Six session in total (medium-term duration of intervention) Massage provider: mothers with trained professional supervision in classes (International Association of Infant Massage)
Outcomes	Types of outcome: Infant Characteristic Questionnaire (ICQ), Global ratings for mother-infant interactions attachment patterns (Strange Situation Procedure) and distractibility Maternal outcomes also reported: Depressive symptoms (EPDS), anxiety (SSAI), bonding scores at 1 year, a baby care questionnaire Timing of assessment: baseline (9 to 12 weeks of infant age) 19 weeks (infant age), and one year follow-up
Notes	Funder: The Foundation for integrated Health.

<i>Risk of bias</i>		
<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Quote p.190 "prospective block-controlled randomised design" Comment: probably done. From investigator "...by block as we needed to ensure that there were sufficient mothers in the support group at any one time (with pure randomisation, we risked having only 1 person in the support "group" or having too many people. So, mothers were contacted and invited to take part in either the massage group OR the support group depending on which arm we were recruiting for at that given timepoint."
Allocation concealment (selection bias)	Unclear risk	No details given.
Incomplete outcome data (attrition bias) All outcomes	Low risk	14/45 did not complete massage group; 20/50 did not complete the support group, no statistical differences between the groups "A Chi-square analysis was conducted to investigate differences between the massage and support group in the number of drop-outs and the numbers who completed all measures at one year, questionnaire measures only at one year or no measures at all at one year. No significant difference was found between the groups (Pearson's Chi square=5.4, ns)." data obtained from study investigator
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote p. 190 "The interactions were rated using the Global Ratings for Mother-Infant Interactions by a blinded, trained rater." ICG was completed by mothers therefore there is a possibility of introducing bias

**Onozawa 2001**

Methods	Design: randomised controlled trial. Setting: community parenting class, UK.
Participants	34 primiparous depressed mothers and their infants aged 9 weeks. Intervention n = 19; control n = 15
Interventions	Infant massage for 1 hour weekly over 5 weeks, plus support group for both intervention and control mothers (medium-term duration of intervention) Massage provider: mothers trained by researchers.
Outcomes	1. EPDS 2. Assessment of mother-infant interaction
Notes	Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	Unclear risk	Described as randomised, no details given.
Allocation concealment (selection bias)	Unclear risk	Unclear.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	7/19 (Intervention); 2/15 (control) did not complete, mainly due to inconvenient time of sessions. Dropouts not evenly distributed between the groups 35% of the sample dropped out because the time of the class was inconvenient (7 from the massage and 2 from the control group did not complete and a further 2 mothers in the massage group and 1 in the control group did not have interactions recorded because their infants were unsettled) Infants who started and did not complete the study were not significantly different demographically from those that completed Comment: judged as unclear risk of bias as the dropouts were not evenly distributed between the groups
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.

**Onozawa 2001** (Continued)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	The assessment of mother-infant interaction scores was completed by the researcher who was aware of the infants' allocation groups, but 10 dyads were coded by an experienced independent rater who was blind to study group and the researcher's reliability ratings were checked against the blinded coder. Two groups of dimensions did not meet the reliability standards and these were eliminated from the study
---	--------------	--

**Oswalt 2007**

Methods	Design: randomised controlled trial. Setting: school-based parent training programme for adolescent mothers, USA
Participants	Sample sizes: n = 21; intervention n = 9; control n = 16. Ages: intervention 52.71days (SD 24.18); control 84.00 days (SD 64.67) Gender: not stated.
Interventions	Infants massaged daily for approximately 30 min, by mothers trained in massage daily for 2 months (medium-term duration of intervention). Also enrolled in parent training. Mothers were teenagers Massage provider: mothers.
Outcomes	Types of outcome: infant PSI child domain, weight/growth scores requested Maternal: PSI parent domain, MCQ and BDI, a non-validated physical contact score also reported Timing of assessment: after 2 months of intervention.
Notes	Funder: unclear - Young Mothers' program?

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote p 285 "using a random number table".
Allocation concealment (selection bias)	Low risk	Sealed envelope (further details from trial investigator).
Incomplete outcome data (attrition bias) All outcomes	High risk	7/9 intervention; 8/16 control completed. Data for only 15 of 25 participants was obtained due to difficulty in tracking participants. Analysis based on these 15 completers



**Oswalt 2007** (Continued)

Selective reporting (reporting bias)	Unclear risk	All pre-specified outcomes reported, except that mothers were asked to complete a worksheet but no worksheets were completed and returned
Blinding of participants and personnel (performance bias) All outcomes	High risk	No details given but unlikely given the nature of the intervention
Blinding of outcome assessment (detection bias) All outcomes	High risk	Outcomes assessors were not blind to the allocation (further details from trial investigator)

**Shao 2005**

Methods	Design: quasi-randomised controlled trial. Setting: unclear, China.
Participants	210 newborn infants: n = 105 intervention; n = 105 control.
Interventions	15 minutes of massage twice daily over 30 days versus a 'no treatment' control group (short duration of intervention) Massage provider: unclear
Outcomes	Weight.
Notes	Paper not fully transcribed. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Quasi-randomised according to sequence of birth time.
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	High risk. Unclear which measurements and how measurements were taken

**Shao 2005** (Continued)

		Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Shi 2002**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	80 newborn infants; n = 40 intervention; n = 40 control.
Interventions	15 minutes of massage twice daily over 28 days versus a 'no treatment' control group (short duration of intervention) Massage provider: unclear.
Outcomes	Weight and height.
Notes	Paper not fully transcribed. Funder: unclear.

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given. Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	High risk. Unclear how measurements were taken. Comment: judged as high risk, no further details available from trial investigator

**Shi 2002** (Continued)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Sun 2004**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	210 newborn infants: n = 105 intervention; n = 105 control.
Interventions	15 minutes of massage twice daily over 42 days versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: unclear.
Outcomes	Weight, bilirubin and sleeping time.
Notes	Paper not fully transcribed. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No details given, no apparent attempt to conceal allocation. Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	High risk. Unclear how measurements were taken. Comment: judged as high risk, no further details available from trial investigator

**Sun 2004** (Continued)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Wang 1999**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	60 newborn infants: n = 30 intervention; n = 30 control.
Interventions	15 minutes of massage three times daily over 42 days versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: unclear
Outcomes	Weight.
Notes	Paper not fully transcribed. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	High risk. Unclear how measurements were taken. Comment: judged as high risk, no further details available from trial investigator

**Wang 1999** (Continued)

Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator

**Wang 2001**

Methods	Design: randomised controlled trial. Setting: maternity ward then at home (community).
Participants	Sample sizes: n = 57; intervention n = 27; control n = 30. Ages: commenced within 24 hours of birth. Gender: not stated.
Interventions	Duration, dose, type. 15-20 min per day started by trained professionals continued daily by the mother after discharge for 2 months (medium-term duration of intervention). 30 days follow-up medical staff to check massage technique telephone numbers provided in case of problem Control group received no massage (routine care only). Massage provider: trained professionals then mothers.
Outcomes	Types of outcome: 0-3 education development checklist (Capital Institute of Children 0-3 years old checklist), weight Timing of assessment: 60 days for development checklist, 2 months for weight
Notes	Funder: not stated.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Described as randomly divided, no further details.
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	Results for n = 57 are reported (number included in study). No dropouts reported. Dropouts or losses to follow-up not addressed in the study report

**Wang 2001** (Continued)

Selective reporting (reporting bias)	Low risk	All outcomes reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Blind outcome assessment (0-3 development checklist) stated but unclear who is blinded and how

**White-Traut 2009**

Methods	Design: quasi-randomised controlled trial. First participant assigned a random number, followed by alternate allocation of subsequent participants Setting: maternity hospital, USA.
Participants	Sample sizes: n = 40: intervention ATVV n = 16; control n = 10; (tactile only group n = 14) Ages: intervention 36.32 hours (SD 10.50); control 34.29 hours (SD 7.18) Gender: intervention 62.5% male; control 30% male.
Interventions	Duration, dose, type. Infants were randomly assigned to receive one 15 minute session of tactile-only, versus auditory, tactile, visual, and vestibular, versus no stimulation 30 minutes before feeding (brief duration of intervention) Massage provider: researchers.
Outcomes	Types of outcome: salivary cortisol ( $\mu\text{g/dL}$ ) and behavioural state Timing of assessment: salivary cortisol baseline, immediately post-intervention and 10 min post-intervention Behavioural state at baseline, mid intervention post-intervention (Thoman 1987).
Notes	Funder: the Harris Foundation.

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote p. 27 "via random start in a random numbers table". Note from investigator: "We used a random start in a random numbers table. The control group was selected when the next even number and the experimental group was selected when the next number was an odd number." Comment: quasi-randomised.

Allocation concealment (selection bias)	Low risk	Study investigator describes allocation by sealed envelopes in a personal communication
Incomplete outcome data (attrition bias) All outcomes	Low risk	40/60 contributed to the final salivary cortisol analyses, due to insufficient sample volumes being collected Comment: sample size was not evenly distributed between the groups at different time points Behavioural state is only reported for the same numbers of infants for each group and time point as were available for cortisol analysis Participants for who complete data only was available were analysed
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes reported. Behavioural analysis: no infants were observed in the indeterminate state category and it was dropped from the analysis. Comment: we considered that it is unlikely that this finding could bias the results of the study
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: unlikely given the nature of the intervention, risk of bias is likely to be high
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote and full details p. 27 “behavioural state was judged by a research assistant who was blinded to group assignment” Study investigator reports “Blinding of participants and personnel (performance bias) mothers were told which group the baby was assigned to mothers did not observe the protocol. The person judging state wore a head set and turned away while the intervention was conducted. They remained blinded to group assignment because the intervention was stopped while they coded behavioral state” (personal communication)

**Xua 2004**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	124 newborn infants; n = 61 intervention; n = 63 control.
Interventions	15 - 20 minutes of massage twice daily over three months versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: unclear.
Outcomes	Duration of sleep; frequency of night wakes and crying; length of crying; length of time for normal sleeping pattern
Notes	Paper not fully transcribed. Also present results for six-month follow-up. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	Duration of sleep, frequency of night wakes and crying (length of crying) were prespecified outcomes The paper also included the incidence of sleep disturbances and length of time required to develop a normal sleeping pattern but the significance of these results was not explored in the paper Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.



**Xua 2004** (Continued)

Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
---	-----------	---

**Ye 2004**

Methods	Design: randomised controlled trial. Setting: unclear, China.
Participants	100 newborn infants: n = 50 intervention; n = 50 control.
Interventions	10 - 15 minutes of massage twice daily over 42 days versus a 'no treatment' control group (medium-term duration of intervention) Massage provider: unclear.
Outcomes	Weight.
Notes	Paper not fully transcribed. Funder: unclear.

***Risk of bias***

<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	High risk	Measurement methods unclear. Comment: judged as high risk, no further details available from trial investigator
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.
Blinding of outcome assessment (detection bias)	Low risk	Assessors were blinded.

All outcomes		
<b>Zhai 2001</b>		
Methods	Design: quasi-randomised controlled trial. Setting: unclear, China.	
Participants	100 newborn infants: n = 50 intervention; n = 50 control.	
Interventions	15 minutes of massage three times daily over 30 days versus a 'no treatment' control group (short duration of intervention) Massage provider: unclear.	
Outcomes	Weight, height and head circumference, means of change scores only, no SD	
Notes	Paper not fully transcribed. Funder: unclear.	
<b>Risk of bias</b>		
<b>Bias</b>	<b>Authors' judgement</b>	<b>Support for judgement</b>
Random sequence generation (selection bias)	High risk	Quasi-randomised according to sequence of admission number: even numbers assigned to massage group. Odd numbers assigned to control group
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	Unclear risk	All pre-specified physical growth outcomes are apparently reported, but milk intake was also reported Comment: we judged this as unclear as it is not clear if reporting of additional outcome measures could bias the study results
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention.

**Zhai 2001** (Continued)

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome assessors blinded.
---	----------	----------------------------

**Zhu 2010**

Methods	Design: quasi-randomised controlled trial (assigned to treatment or control on basis of odd and even days of birth) Setting: community (initiated in hospital). China.
Participants	Sample sizes: n = 115; intervention n = 55; control n = 60. Ages: neonates, not otherwise specified. Gender: intervention 45% male; control 45% male.
Interventions	15-20 min per session, 2-3 times a day for 3 months (medium-term duration of intervention). Care as usual in control group Massage provider: parents.
Outcomes	Types of outcome: 1 month after birth neonatal behavioral neurological assessment score (NBNA); at 3 months, using adapted China Institute of Psychology and Child Development Center scales mental development index (MDI), psychomotor development index (PDI) and head circumference measurements Timing of assessment: 1 month and 3 months.
Notes	Funder: unclear (not stated).

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quasi-randomised by odd and even date of birth
Allocation concealment (selection bias)	High risk	No apparent attempt to conceal allocation, no details given Comment: judged as high risk, no further details available from trial investigator
Incomplete outcome data (attrition bias) All outcomes	High risk	No dropouts reported. Dropouts or losses to follow-up not addressed in the study report
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes are apparently reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible due to nature of intervention

**Zhu 2010** (Continued)

Blinding of outcome assessment (detection bias) All outcomes	High risk	No details given. Comment: judged as high risk, no further details available from trial investigator
---	-----------	---

ATVV = auditory, tactile, visual, vestibular

BDI = Beck Depression Inventory

EPDS = Edinburgh Postnatal Depression Scale

Hb = haemoglobin

MCQ = Maternal Confidence Questionnaire

SD = standard deviation

SSAI = Spielberger State Anxiety Index

URTI = upper respiratory infection

**Duration of intervention**

Brief = a single session

Short = intervention took place for up to 4 weeks

Medium-term = intervention took place for at least 4 weeks and up to 12 weeks

Long = intervention took place for at least 12 weeks and up to 26 weeks

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

Study	Reason for exclusion
Clarke 2000	Trial was not randomised
Cullen 2000	Infants participating in the study were aged between 3 and 14 months (mean 7.1 SD = 3.4)- outside of the stated aged range for this review
Darmstadt 2002a	Large survey
Fernandez 1998	Spanish study of paediatric massage - not a RCT
Field 2000b	Study intervention aimed at mothers rather than infants - consisted of free day care for the infants and a rehab program (social, educational, and vocational) plus several mood induction interventions for the mothers, including relaxation, massage therapy, and mother-infant interaction coaching
Field 2004	Study compared infants who either received light pressure or moderate pressure massage. There was no control group
Fogaça 2005	Not randomised, no control group.
Huhtala 2000	Study compared infant massage and crib vibrator interventions. There was no control group
Im 2007	RCT of Yakson massage versus non-nutritive sucking , versus control. Excluded as this is a study of pain relief in infants (pain due to heel stick test)

(Continued)

Ineson 1995	This was a review article of literature, not a RCT
Jing L 2007	RCT, massage and use of educational toy, control group received no intervention. Infants under 6 months of age. Growth, physical and mental development indices reported. Excluded due to multimodal nature of intervention
Jump 2006	RCT of orphaned infants in Ecuador. Randomly assigned to intervention or control. Outcome was number of days of illness. Excluded as children too old, mean 10.6 month in experimental, 10.4 months in control group
Lee 2006	Not an RCT (non equivalent control group pretest-post test design)
Li 2002	Non randomised. Employed 'convenient sampling' method that is, mothers who volunteered were in the massage group the control group were mother who did not carry out massage. (Further information from trial investigator)
Oswalt 2009	Dissertation, RCT, massage and control group, outcomes included maternal outcomes of stress, depression and confidence. Excluded as participant group was HIV infected mothers
Pardew 1996	This dissertation investigated the effects of infant massage on interactions between high risk infants and their care givers
Park 2006	RCT, control group (no massage) versus Yakson massage. Excluded as this is a study of pain relief in infants (pain due to heel stick test)
Peláez-Nogueras 1996	Measured infant affect during the 'still-face' procedure only
Peláez-Nogueras 1997	This study was a test procedure for measuring eye contact when infants were touched or not touched
Peláez-Nogueras1997b	This study compares stroking with tickling and poking on infant eye contact
Scafidi 1996	Sample comprised HIV-exposed infants with a lower gestational age and birthweight than normal
Serrano 2010	Not an RCT (age matched control group who did not receive massage)
Stack 1990	Measured tactile stimulation during the still face procedure only
Yilmaz 2009	Not an RCT, cross matched control group.
Zhu 2000	Randomised 'randomly selected', normal and sick term and preterm neonates, compares different types of massage, does not include a no treatment control group
Zhu 2002	Randomised experimental animal study.

RCT: randomised controlled trial

SD: standard deviation

## DATA AND ANALYSES

### Comparison 1. Infant massage versus control - physical development

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Weight	18		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 Post-intervention	18	2271	Mean Difference (IV, Random, 95% CI)	-965.25 [-1360.52, -569.98]
1.2 Post-intervention Western studies	2	81	Mean Difference (IV, Random, 95% CI)	-127.10 [-575.14, 320.93]
1.3 Post-intervention sensitivity analysis for Kim 2003	17	2213	Mean Difference (IV, Random, 95% CI)	-975.96 [-1390.63, -561.30]
1.4 Post-intervention sensitivity analysis risk of bias	3	405	Mean Difference (IV, Random, 95% CI)	-203.55 [-443.37, 36.26]
1.5 Follow-up 6 to 8 months	3	202	Mean Difference (IV, Random, 95% CI)	-758.29 [-1364.67, -151.90]
1.6 Follow-up 6 months sensitivity analysis for Kim 2003	2	157	Mean Difference (IV, Random, 95% CI)	-455.07 [-823.80, -86.33]
2 Weight: subgroup analyses (duration of intervention)	18		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 Post-intervention subgroup short term	5	443	Mean Difference (IV, Random, 95% CI)	-374.07 [-654.84, -93.31]
2.2 Post-intervention subgroup medium term	12	1648	Mean Difference (IV, Random, 95% CI)	-1259.19 [-1807.80, -710.58]
2.3 Post-intervention subgroup long term	1	180	Mean Difference (IV, Random, 95% CI)	-500.00 [-811.25, -188.75]
3 Length	11		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.1 Post-intervention	11	1683	Mean Difference (IV, Random, 95% CI)	-1.30 [-1.60, 1.00]
3.2 Post-intervention sensitivity analysis risk of bias	3	405	Mean Difference (IV, Random, 95% CI)	-0.65 [-1.20, -0.11]
3.3 Follow-up 6 months	2	161	Mean Difference (IV, Random, 95% CI)	-1.98 [-4.69, 0.72]
4 Length: subgroup analyses (duration of intervention)	11		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.1 Post-intervention subgroup short duration	5	443	Mean Difference (IV, Random, 95% CI)	-1.00 [-1.54, -0.47]
4.2 Post-intervention subgroup medium-term duration	5	1060	Mean Difference (IV, Random, 95% CI)	-1.51 [-1.76, -1.27]
4.3 Post-intervention subgroup long duration	1	180	Mean Difference (IV, Random, 95% CI)	-1.13 [-1.88, -0.38]
5 Head circumference	10		Mean Difference (IV, Random, 95% CI)	Subtotals only
5.1 Post-intervention	9	1423	Mean Difference (IV, Random, 95% CI)	-0.81 [-1.18, -0.45]

5.2 Post-intervention sensitivity analysis risk of bias	2	225	Mean Difference (IV, Random, 95% CI)	-0.07 [-0.27, 0.12]
5.3 Follow-up 6 months	2	160	Mean Difference (IV, Random, 95% CI)	-2.19 [-3.88, -0.49]
6 Head circumference: subgroup analyses (duration of intervention)	9		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.1 Post-intervention subgroup short	4	363	Mean Difference (IV, Random, 95% CI)	-0.70 [-1.45, 0.05]
6.2 Post-intervention subgroup medium-term	5	1060	Mean Difference (IV, Random, 95% CI)	-0.90 [-1.16, -0.64]
7 Mid arm circumference	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.1 Post-intervention	2	225	Mean Difference (IV, Random, 95% CI)	-0.47 [-0.80, -0.13]
8 Mid leg/thigh circumference	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.1 Post-intervention	2	225	Mean Difference (IV, Random, 95% CI)	-0.31 [-0.49, -0.13]
9 Abdominal circumference	1	100	Mean Difference (IV, Random, 95% CI)	-0.75 [-1.09, -0.41]
9.1 Post-intervention	1	100	Mean Difference (IV, Random, 95% CI)	-0.75 [-1.09, -0.41]
10 Chest circumference	1	100	Mean Difference (IV, Random, 95% CI)	-0.88 [-1.22, -0.54]
10.1 Post-intervention	1	100	Mean Difference (IV, Random, 95% CI)	-0.88 [-1.22, -0.54]
11 Hormones: cortisol	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
11.1 Salivary cortisol immediately post-intervention	1	19	Std. Mean Difference (IV, Random, 95% CI)	0.46 [-0.45, 1.38]
11.2 Salivary cortisol - 10 to 20 min post-intervention	2	54	Std. Mean Difference (IV, Random, 95% CI)	-0.24 [-0.77, 0.30]
11.3 Urinary cortisol - day 12 of intervention	1	40	Std. Mean Difference (IV, Random, 95% CI)	-0.80 [-1.45, -0.15]
12 Hormones: norepinephrine	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
12.1 Post-intervention	1	40	Mean Difference (IV, Random, 95% CI)	-60.3 [-111.88, -8.72]
13 Hormones: epinephrine	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
13.1 Post-intervention	1	40	Mean Difference (IV, Random, 95% CI)	-13.00 [-20.08, -5.92]
14 Hormones: serotonin	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
14.1 Post-intervention	1	40	Mean Difference (IV, Random, 95% CI)	-295.5 [-705.25, 114.25]
15 Hormones: 6-sulphatoxymelatonin secretion	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
16 Biochemical markers: Bilirubin (7 days PN)	2	410	Mean Difference (IV, Random, 95% CI)	-38.11 [-50.61, -25.61]
17 Crying or fussing time	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
17.1 Post-intervention	4	341	Mean Difference (IV, Random, 95% CI)	-0.36 [-0.52, -0.19]
17.2 Follow-up 3 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.21 [-0.40, -0.02]
17.3 Follow-up 6 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.15 [-0.29, -0.01]
18 Crying frequency (times)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
18.1 Post-intervention	1	124	Mean Difference (IV, Random, 95% CI)	-0.34 [-0.56, -0.12]
18.2 Follow-up 3 months	1	126	Mean Difference (IV, Random, 95% CI)	-0.19 [-0.36, -0.02]
18.3 Follow-up 6 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.18 [-0.35, -0.01]
19 Sleep/wake behaviours (Thoman)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
19.1 Quiet sleep	1	40	Mean Difference (IV, Random, 95% CI)	-6.30 [-20.16, 7.56]
19.2 Active sleep	1	40	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

19.3 Inactive alert	1	40	Mean Difference (IV, Random, 95% CI)	-12.70 [-19.38, -6.02]
19.4 Crying	1	40	Mean Difference (IV, Random, 95% CI)	-8.2 [-12.24, -4.16]
19.5 Drowsy	1	40	Mean Difference (IV, Random, 95% CI)	2.0 [-0.19, 4.19]
19.6 Active awake	1	40	Mean Difference (IV, Random, 95% CI)	-15.00 [-22.29, -7.71]
19.7 REM sleep	1	40	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
19.8 Movement	1	40	Mean Difference (IV, Random, 95% CI)	-12.60 [-27.59, 2.39]
20 Behavioural state immediately post-intervention (Thoman)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
20.1 Asleep	1	26	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.55, 1.96]
20.2 Awake	1	26	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.27, 2.23]
20.3 Crying	1	26	Risk Ratio (M-H, Random, 95% CI)	1.94 [0.09, 43.50]
21 Sleep duration over 24hr period	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
21.1 Post-intervention	4	634	Mean Difference (IV, Random, 95% CI)	-0.91 [-1.51, -0.30]
21.2 Sleep follow-up 3 months	1	124	Mean Difference (IV, Random, 95% CI)	-1.30 [-1.81, -0.79]
21.3 Sleep follow-up 6 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.08 [-0.64, 0.48]
22 Mean increase in 24h sleep	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
22.1 Post-intervention	2	225	Std. Mean Difference (IV, Random, 95% CI)	-1.47 [-4.43, 1.49]
23 Mean increase in duration of night sleep	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
23.1 Post-intervention	2	225	Std. Mean Difference (IV, Random, 95% CI)	-1.28 [-3.66, 1.10]
24 Mean increase in duration of day sleep	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
24.1 Post-intervention	1	125	Mean Difference (IV, Random, 95% CI)	0.10 [-0.21, 0.41]
25 Mean increase in duration of first morning sleep after massage	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
25.1 Post-intervention	1	125	Mean Difference (IV, Random, 95% CI)	-1.52 [-1.69, -1.35]
26 Sleep (total hours per night)	1	100	Mean Difference (IV, Random, 95% CI)	-0.70 [1.00, -0.40]
26.1 Post-intervention	1	100	Mean Difference (IV, Random, 95% CI)	-0.70 [1.00, -0.40]
27 Number of naps (total number of naps)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
28 Number of naps in day	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
29 Number of naps at night	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
30 Night Wake Frequency (times)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
30.1 Post-intervention	1	124	Mean Difference (IV, Random, 95% CI)	-0.48 [-0.81, -0.15]
30.2 Follow-up 3 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.38 [-0.63, -0.13]
30.3 Follow-up 6 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.35 [-0.56, -0.14]
31 Night wake duration	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
31.1 Post-intervention	1	124	Mean Difference (IV, Random, 95% CI)	-0.27 [-0.51, -0.03]
31.2 Follow-up 3 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.18 [-0.31, -0.05]
31.3 Follow-up 6 months	1	124	Mean Difference (IV, Random, 95% CI)	-0.26 [-0.50, -0.02]
32 Blood flow (post intervention)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
32.1 Blood flow (cm/s) post-intervention	1	125	Mean Difference (IV, Random, 95% CI)	-0.54 [-1.03, -0.05]
32.2 Blood velocity (cm/s) post-intervention	1	125	Mean Difference (IV, Random, 95% CI)	-0.98 [-6.65, 4.69]



32.3 Vessel diameter (cm) post-intervention	1	125	Mean Difference (IV, Random, 95% CI)	0.02 [0.01, 0.03]
33 Formula intake	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
33.1 Post-intervention (US fl oz converted to ml)	1	40	Mean Difference (IV, Random, 95% CI)	70.97 [6.16, 135.78]
34 Illness	2		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
34.1 URTI (post intervention)	2	310	Risk Ratio (M-H, Random, 95% CI)	1.19 [0.86, 1.65]
34.2 Anaemia (post intervention)	2	310	Risk Ratio (M-H, Random, 95% CI)	1.49 [0.79, 2.82]
34.3 Diarrhoea (post intervention)	2	310	Risk Ratio (M-H, Random, 95% CI)	0.39 [0.20, 0.76]
35 Illness and clinic visits	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
35.1 Illness follow-up 6 months	1	45	Mean Difference (IV, Random, 95% CI)	-8.82 [-10.62, -7.02]
35.2 Clinic visits follow-up 6 months	1	45	Mean Difference (IV, Random, 95% CI)	-5.98 [-7.07, -4.89]

## Comparison 2. Infant massage versus control - mental health and development

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Infant temperament meta-analyses	3		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 Activity (post-intervention)	3	121	Std. Mean Difference (IV, Random, 95% CI)	0.39 [-0.34, 1.13]
1.2 Persistence (post-intervention)	2	81	Std. Mean Difference (IV, Random, 95% CI)	0.24 [-0.20, 0.67]
1.3 Soothability (post-intervention)	2	80	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.94, 0.35]
2 Infant temperament (CCTI) post intervention	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 Activity	1	40	Mean Difference (IV, Random, 95% CI)	-1.60 [-4.41, 1.21]
2.2 Soothability	1	40	Mean Difference (IV, Random, 95% CI)	-2.90 [-5.71, -0.09]
2.3 Emotionality	1	40	Mean Difference (IV, Random, 95% CI)	-0.80 [-3.61, 2.01]
2.4 Sociability	1	40	Mean Difference (IV, Random, 95% CI)	-1.5 [-3.98, 0.98]
2.5 Persistence	1	40	Mean Difference (IV, Random, 95% CI)	0.10 [-2.38, 2.58]
2.6 Food adaptation	1	40	Mean Difference (IV, Random, 95% CI)	0.5 [-1.98, 2.98]
3 Infant temperament (Infant behaviour questionnaire (IBQ) post intervention)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.1 Activity	1	40	Mean Difference (IV, Random, 95% CI)	0.56 [0.08, 1.04]
3.2 Soothability	1	40	Mean Difference (IV, Random, 95% CI)	0.03 [-0.59, 0.65]
3.3 Duration of orienting	1	40	Mean Difference (IV, Random, 95% CI)	0.0 [-0.82, 0.82]
3.4 Distress to limitations	1	40	Mean Difference (IV, Random, 95% CI)	-0.08 [-0.49, 0.33]
3.5 Fear	1	40	Mean Difference (IV, Random, 95% CI)	-0.06 [-0.63, 0.51]
3.6 Amount of smiling	1	40	Mean Difference (IV, Random, 95% CI)	0.30 [-0.14, 0.74]

4 Infant temperament questionnaire (revised RITQ (Carey)) post-intervention 4 months	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.1 Activity	1	41	Mean Difference (IV, Random, 95% CI)	0.41 [0.11, 0.71]
4.2 Rhythmicity	1	41	Mean Difference (IV, Random, 95% CI)	-0.19 [-0.63, 0.25]
4.3 Approach	1	41	Mean Difference (IV, Random, 95% CI)	0.17 [-0.18, 0.52]
4.4 Adaptability	1	41	Mean Difference (IV, Random, 95% CI)	0.10 [-0.30, 0.50]
4.5 Intensity	1	41	Mean Difference (IV, Random, 95% CI)	0.19 [-0.28, 0.66]
4.6 Mood	1	41	Mean Difference (IV, Random, 95% CI)	0.31 [-0.14, 0.76]
4.7 Persistence	1	41	Mean Difference (IV, Random, 95% CI)	0.33 [-0.11, 0.77]
4.8 Distractibility	1	41	Mean Difference (IV, Random, 95% CI)	0.28 [-0.18, 0.74]
4.9 Threshold	1	41	Mean Difference (IV, Random, 95% CI)	0.11 [-0.43, 0.65]
5 Infant temperament questionnaire (revised RITQ (Carey)) follow-up 8 months	1	369	Mean Difference (IV, Random, 95% CI)	0.66 [0.48, 0.84]
5.1 Activity	1	41	Mean Difference (IV, Random, 95% CI)	0.25 [-0.33, 0.83]
5.2 Rhythmicity	1	41	Mean Difference (IV, Random, 95% CI)	0.80 [0.12, 1.48]
5.3 Approach	1	41	Mean Difference (IV, Random, 95% CI)	0.88 [0.25, 1.51]
5.4 Adaptability	1	41	Mean Difference (IV, Random, 95% CI)	0.69 [0.01, 1.37]
5.5 Intensity	1	41	Mean Difference (IV, Random, 95% CI)	0.39 [0.02, 0.76]
5.6 Mood	1	41	Mean Difference (IV, Random, 95% CI)	1.08 [0.65, 1.51]
5.7 Persistence	1	41	Mean Difference (IV, Random, 95% CI)	0.65 [-0.03, 1.33]
5.8 Distractibility	1	41	Mean Difference (IV, Random, 95% CI)	0.72 [0.32, 1.12]
5.9 Threshold	1	41	Mean Difference (IV, Random, 95% CI)	0.48 [-0.27, 1.23]
6 Infant Care Questionnaire post-intervention	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
6.1 ICQ fussy/difficult	1	59	Mean Difference (IV, Random, 95% CI)	1.37 [-2.53, 5.27]
6.2 ICQ unadaptable	1	59	Mean Difference (IV, Random, 95% CI)	-0.19 [-1.51, 1.13]
6.3 ICQ dull	1	59	Mean Difference (IV, Random, 95% CI)	-1.08 [-2.60, 0.44]
6.4 ICQ unpredictable	1	59	Mean Difference (IV, Random, 95% CI)	0.61 [-1.78, 3.00]
7 Infant Care Questionnaire follow-up 1 year	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
7.1 ICQ fussy/difficult	1	50	Mean Difference (IV, Random, 95% CI)	1.05 [-2.43, 4.53]
7.2 ICQ unadaptable	1	50	Mean Difference (IV, Random, 95% CI)	-0.39 [-1.63, 0.85]
7.3 ICQ dull	1	50	Mean Difference (IV, Random, 95% CI)	0.35 [-1.54, 2.24]
7.4 ICQ unpredictable	1	50	Mean Difference (IV, Random, 95% CI)	1.89 [-0.55, 4.33]
8 Infant attachment (Q set)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
8.1 Follow-up 1 year	1	39	Mean Difference (IV, Random, 95% CI)	-0.06 [-0.17, 0.05]
9 Child behaviour (HOME)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
9.1 Follow-up (24 months)	1	25	Mean Difference (IV, Random, 95% CI)	0.34 [-1.92, 2.60]
10 Eyberg Child Behaviour Inventory (ECBI) - Intensity domain	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
10.1 Follow-up 24 months	1	25	Mean Difference (IV, Random, 95% CI)	4.95 [-9.94, 19.84]
11 Eyberg Child Behaviour Inventory (ECBI) - Problem domain	1	25	Mean Difference (IV, Random, 95% CI)	-0.19 [-3.26, 2.88]
11.1 Follow-up 24 months	1	25	Mean Difference (IV, Random, 95% CI)	-0.19 [-3.26, 2.88]
12 Mother and child interaction meta-analysis - Total NCATS and Murray Global	4		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only

12.1 Post-intervention	3	131	Std. Mean Difference (IV, Random, 95% CI)	-0.26 [-1.01, 0.48]
12.2 Follow-up 12 and 24 months	2	65	Std. Mean Difference (IV, Random, 95% CI)	-0.20 [-0.69, 0.29]
13 Nursing Child Feeding Assessment Scale (NCAFS) - Total	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
13.1 Post-intervention (16 weeks)	1	47	Mean Difference (IV, Random, 95% CI)	-2.10 [-6.16, 1.96]
14 Nursing Child Assessment Teaching Scale (NCATS) - Mother	1		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
14.1 Follow-up 24 months	1	25	Std. Mean Difference (IV, Random, 95% CI)	-0.18 [-0.96, 0.61]
15 Nursing Child Assessment Teaching Scale (NCATS) - Child	1	25	Std. Mean Difference (IV, Random, 95% CI)	0.35 [-0.44, 1.14]
15.1 Follow-up 24 months	1	25	Std. Mean Difference (IV, Random, 95% CI)	0.35 [-0.44, 1.14]
16 Maternal sensitivity - warm to cold (Murray)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
16.1 Post-intervention	2	84	Mean Difference (IV, Random, 95% CI)	-0.34 [-1.07, 0.40]
16.2 Follow-up 1 year	1	40	Mean Difference (IV, Random, 95% CI)	-0.84 [-1.07, -0.61]
17 Maternal sensitivity - non-intrusive to intrusive (Murray)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
17.1 Post-intervention	2	84	Mean Difference (IV, Random, 95% CI)	-0.10 [-0.85, 0.66]
17.2 Follow-up 1 year	1	40	Mean Difference (IV, Random, 95% CI)	-0.01 [-0.30, 0.28]
18 Maternal sensitivity - remoteness (Murray)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
18.1 Post-intervention	1	40	Mean Difference (IV, Random, 95% CI)	0.08 [-0.32, 0.48]
18.2 Follow-up	1	62	Mean Difference (IV, Random, 95% CI)	-0.14 [-0.40, 0.12]
19 Infant interactions - infant performance - attentive to non attentive (Murray)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
19.1 Post-intervention	2	84	Mean Difference (IV, Random, 95% CI)	-0.47 [-1.47, 0.52]
19.2 Follow-up 1 year	1	40	Mean Difference (IV, Random, 95% CI)	0.18 [-0.18, 0.54]
20 Infant interactions - lively to inert (Murray)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
20.1 Post-intervention	2	84	Mean Difference (IV, Random, 95% CI)	-0.46 [-1.45, 0.53]
20.2 Follow-up 1 year	1	40	Mean Difference (IV, Random, 95% CI)	-0.11 [-0.31, 0.09]
21 Infant interactions - happy to distressed (Murray)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
21.1 Post intervention	2	84	Mean Difference (IV, Random, 95% CI)	-0.35 [-1.29, 0.59]
21.2 Follow-up 1 year	1	40	Mean Difference (IV, Random, 95% CI)	-0.02 [-0.26, 0.22]
22 Parenting stress (PSI Abidin) child characteristics subscale	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
22.1 Post-intervention	2	55	Mean Difference (IV, Random, 95% CI)	-10.85 [-53.86, 32.16]
23 Psychomotor Development Indices (PDI) meta-analysis post-intervention	4		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
23.1 Post-intervention	4	466	Std. Mean Difference (IV, Random, 95% CI)	-0.35 [-0.54, -0.15]

23.2 Post-intervention sensitivity analysis Western studies	1	41	Std. Mean Difference (IV, Random, 95% CI)	0.00 [-0.61, 0.62]
24 Bayley Psychomotor Development Index (PDI) follow-up	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
24.1 Follow-up 8 months	1	41	Mean Difference (IV, Random, 95% CI)	-0.78 [-11.89, 10.33]
24.2 Follow-up 24 months	1	41	Mean Difference (IV, Random, 95% CI)	-7.52 [-16.53, 1.49]
25 Mental Development Indices (MDI) meta-analysis post-intervention	4		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
25.1 Post-intervention	4	466	Std. Mean Difference (IV, Random, 95% CI)	-0.27 [-0.64, 0.11]
25.2 Post-intervention sensitivity analysis Western studies	1	41	Std. Mean Difference (IV, Random, 95% CI)	0.38 [-0.23, 1.00]
26 Bayley Mental Development Index (MDI) follow-up	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
26.1 Follow-up 8 months	1	41	Mean Difference (IV, Random, 95% CI)	22.85 [4.26, 41.44]
26.2 Follow-up 24 months	1	41	Mean Difference (IV, Random, 95% CI)	-8.59 [-18.80, 1.62]
27 Gessel/Capital meta-analysis (post intervention)	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
27.1 Gross motor	2	237	Std. Mean Difference (IV, Random, 95% CI)	-0.44 [-0.70, -0.18]
27.2 Fine motor	2	237	Std. Mean Difference (IV, Random, 95% CI)	-0.61 [-0.87, -0.35]
27.3 Language	2	237	Std. Mean Difference (IV, Random, 95% CI)	-0.82 [-1.67, 0.03]
27.4 Personal-social behaviour	2	237	Std. Mean Difference (IV, Random, 95% CI)	-0.90 [-1.61, -0.18]
28 Gessel Developmental Quotient (post intervention)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
28.1 Adaptive behaviour	1	180	Mean Difference (IV, Random, 95% CI)	-7.07 [-9.75, -4.39]
28.2 Gross motor	1	180	Mean Difference (IV, Random, 95% CI)	-3.97 [-6.99, -0.95]
28.3 Fine motor	1	180	Mean Difference (IV, Random, 95% CI)	-6.89 [-10.18, -3.60]
28.4 Language	1	180	Mean Difference (IV, Random, 95% CI)	-4.15 [-7.03, -1.27]
28.5 Personal-social behaviour	1	180	Mean Difference (IV, Random, 95% CI)	-6.41 [-9.65, -3.17]
29 Capital institute Mental Checklist (post intervention)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
29.1 Gross motor	1	57	Mean Difference (IV, Random, 95% CI)	-0.24 [-0.44, -0.05]
29.2 Fine motor	1	57	Mean Difference (IV, Random, 95% CI)	-0.28 [-0.51, -0.05]
29.3 Cognitive	1	57	Mean Difference (IV, Random, 95% CI)	-0.54 [-0.92, -0.15]
29.4 Language	1	57	Mean Difference (IV, Random, 95% CI)	-0.7 [-0.99, -0.41]
29.5 Social behaviour	1	57	Mean Difference (IV, Random, 95% CI)	-0.70 [-0.97, -0.42]
29.6 IQ	1	57	Mean Difference (IV, Random, 95% CI)	-27.18 [-33.13, -21.23]
30 Gessel Developmental Quotient (follow-up 6 months)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
30.1 Adaptive behaviour	1	116	Mean Difference (IV, Random, 95% CI)	-5.79 [-9.64, -1.94]
30.2 Gross motor	1	116	Mean Difference (IV, Random, 95% CI)	-2.85 [-8.18, 2.48]
30.3 Fine motor	1	116	Mean Difference (IV, Random, 95% CI)	-8.12 [-11.67, -4.57]
30.4 Language	1	116	Mean Difference (IV, Random, 95% CI)	-7.90 [-11.70, -4.10]
30.5 Personal-social behaviour	1	116	Mean Difference (IV, Random, 95% CI)	-6.19 [-9.83, -2.55]
31 Attachment patterns (strange situation procedure)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
31.1 Secure (1 year follow-up)	1	39	Risk Ratio (M-H, Random, 95% CI)	0.82 [0.50, 1.34]

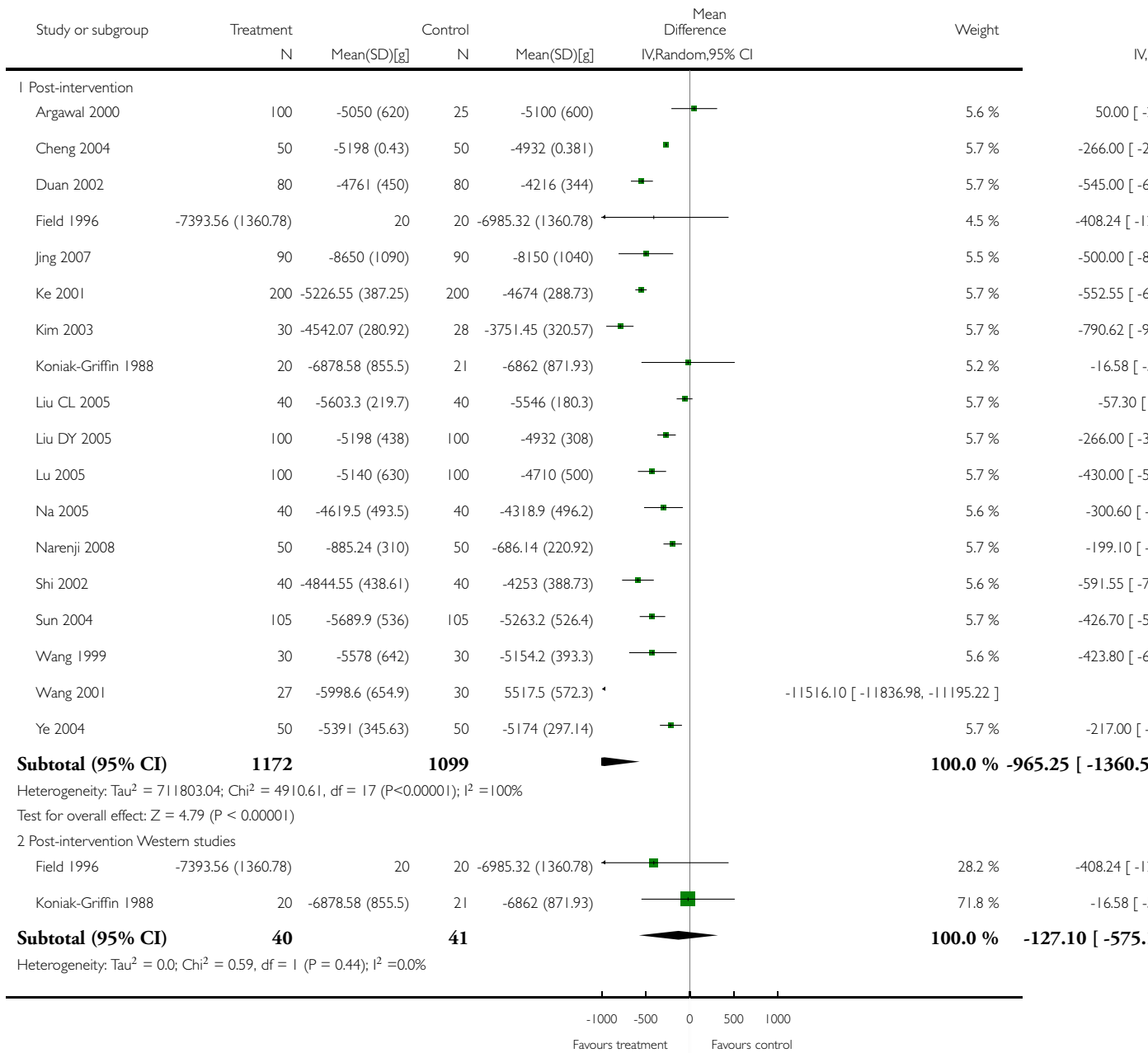
31.2 Avoidant (1 year follow-up)	1	39	Risk Ratio (M-H, Random, 95% CI)	1.39 [0.14, 14.07]
31.3 Resistant (1 year follow-up)	1	39	Risk Ratio (M-H, Random, 95% CI)	3.48 [0.45, 27.02]
31.4 Disorganised (1 year follow-up)	1	39	Risk Ratio (M-H, Random, 95% CI)	0.70 [0.16, 3.02]
32 Distractibility (toy) follow-up 1 year	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
32.1 Mean looks greater than 14 secs	1	32	Risk Ratio (M-H, Random, 95% CI)	2.65 [0.31, 22.82]
32.2 Mean looks less than 14 secs	1	32	Risk Ratio (M-H, Random, 95% CI)	0.88 [0.68, 1.14]
32.3 Max looks greater than 14 secs	1	32	Risk Ratio (M-H, Random, 95% CI)	0.96 [0.66, 1.38]
32.4 Max looks less than 14 secs	1	32	Risk Ratio (M-H, Random, 95% CI)	1.76 [0.37, 8.31]
33 Habituation	1	32	Mean Difference (IV, Random, 95% CI)	-1.10 [-4.79, 2.59]
34 Seconds to habituation	1	32	Mean Difference (IV, Random, 95% CI)	-10.90 [-69.41, 47.61]
35 Trials to habituation	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
36 Post habituation	1	32	Mean Difference (IV, Random, 95% CI)	2.0 [-2.43, 6.43]
37 Habituation test	1	32	Mean Difference (IV, Random, 95% CI)	-12.40 [-19.37, -5.43]

## Analysis 1.1. Comparison 1 Infant massage versus control - physical development, Outcome 1 Weight.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

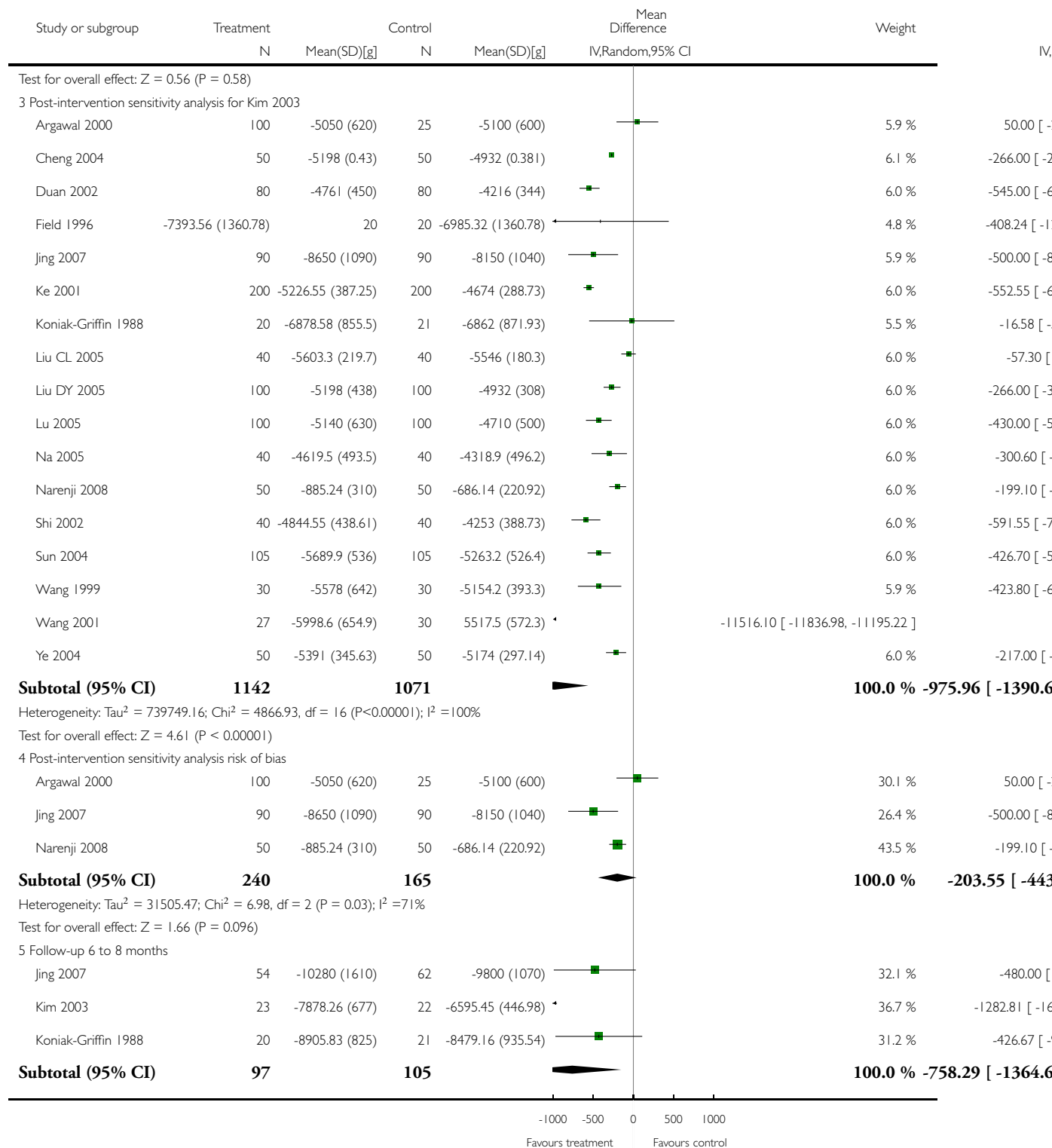
Comparison: 1 Infant massage versus control - physical development

Outcome: 1 Weight



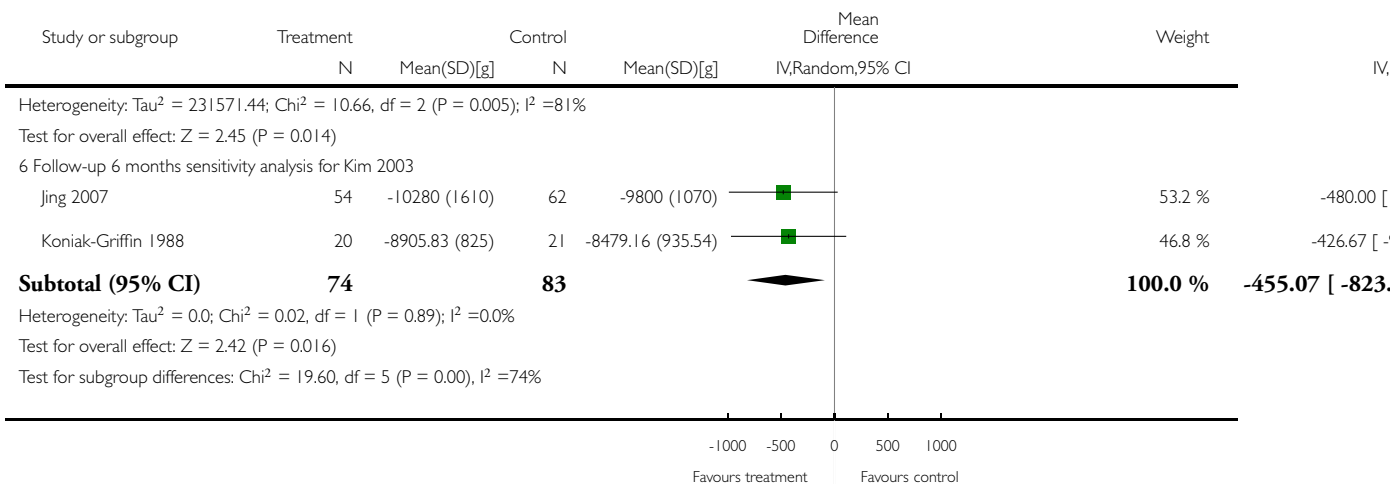
(Continued ...)

(... Continued)



(Continued ...)

(... Continued)



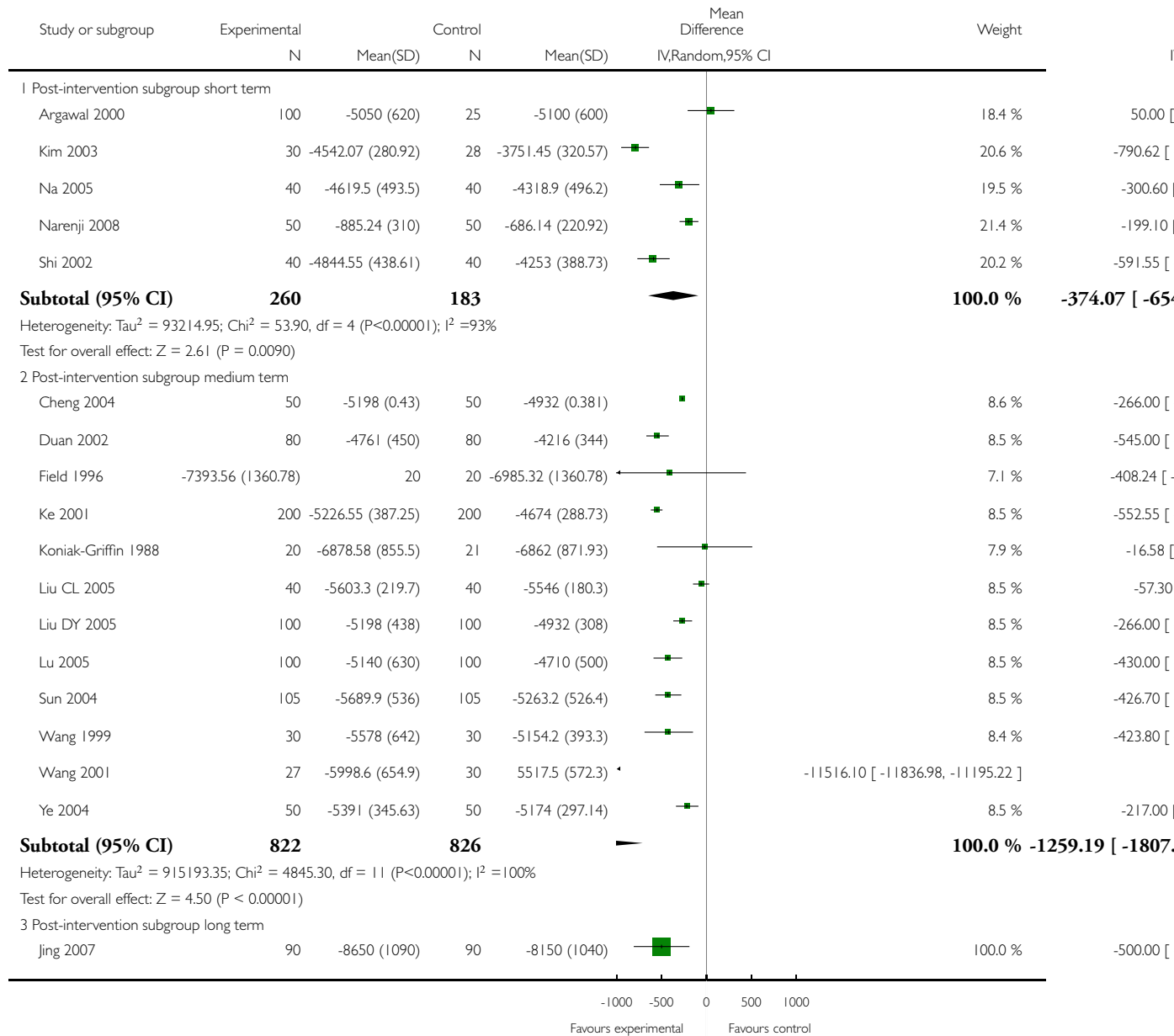


**Analysis 1.2. Comparison 1 Infant massage versus control - physical development, Outcome 2 Weight: subgroup analyses (duration of intervention).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

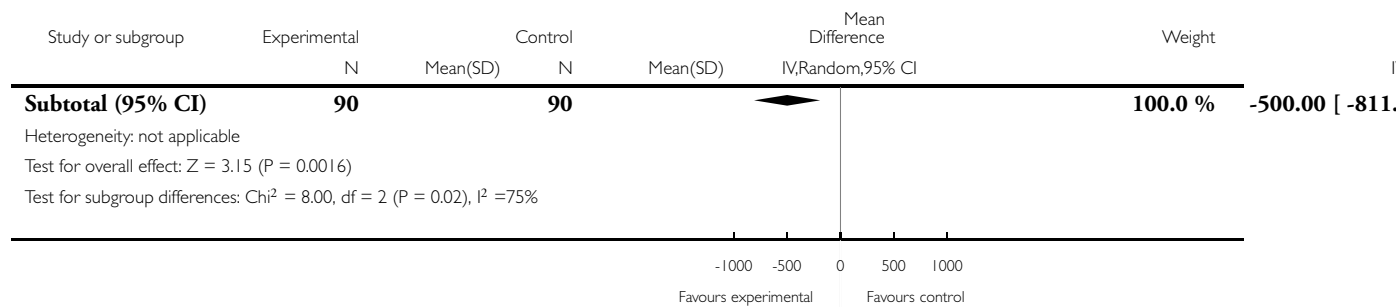
Comparison: 1 Infant massage versus control - physical development

Outcome: 2 Weight: subgroup analyses (duration of intervention)



(Continued ...)

(... Continued)

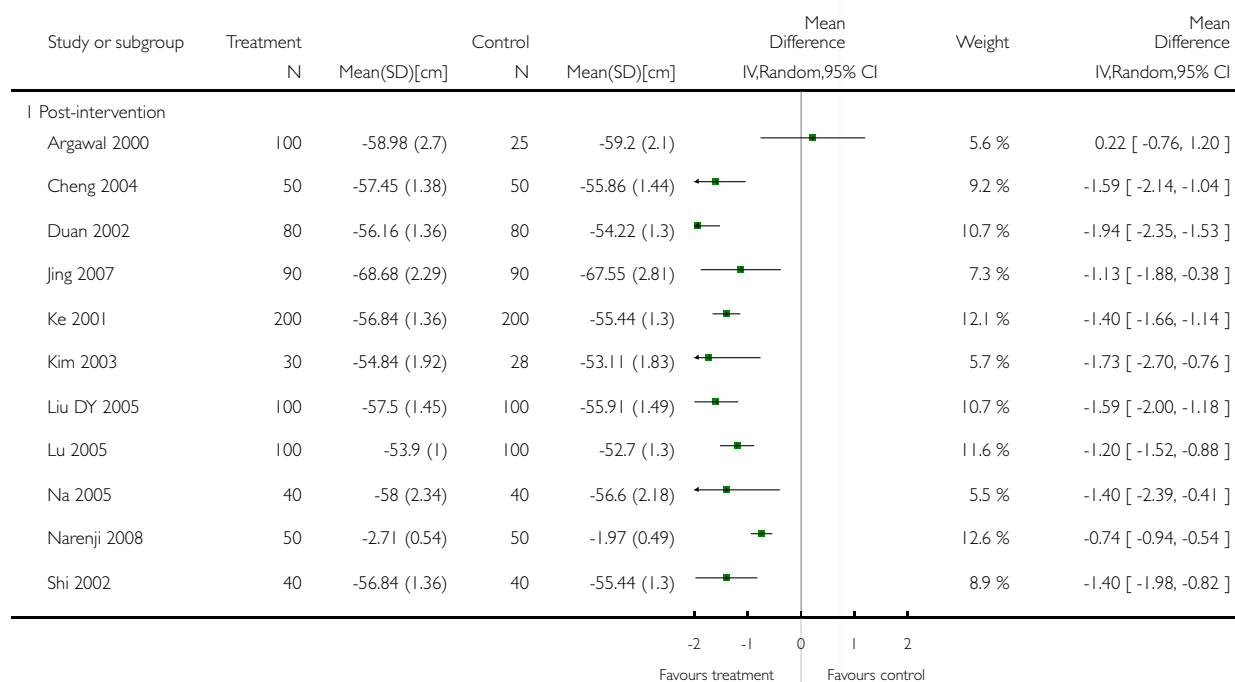


### Analysis 1.3. Comparison 1 Infant massage versus control - physical development, Outcome 3 Length.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

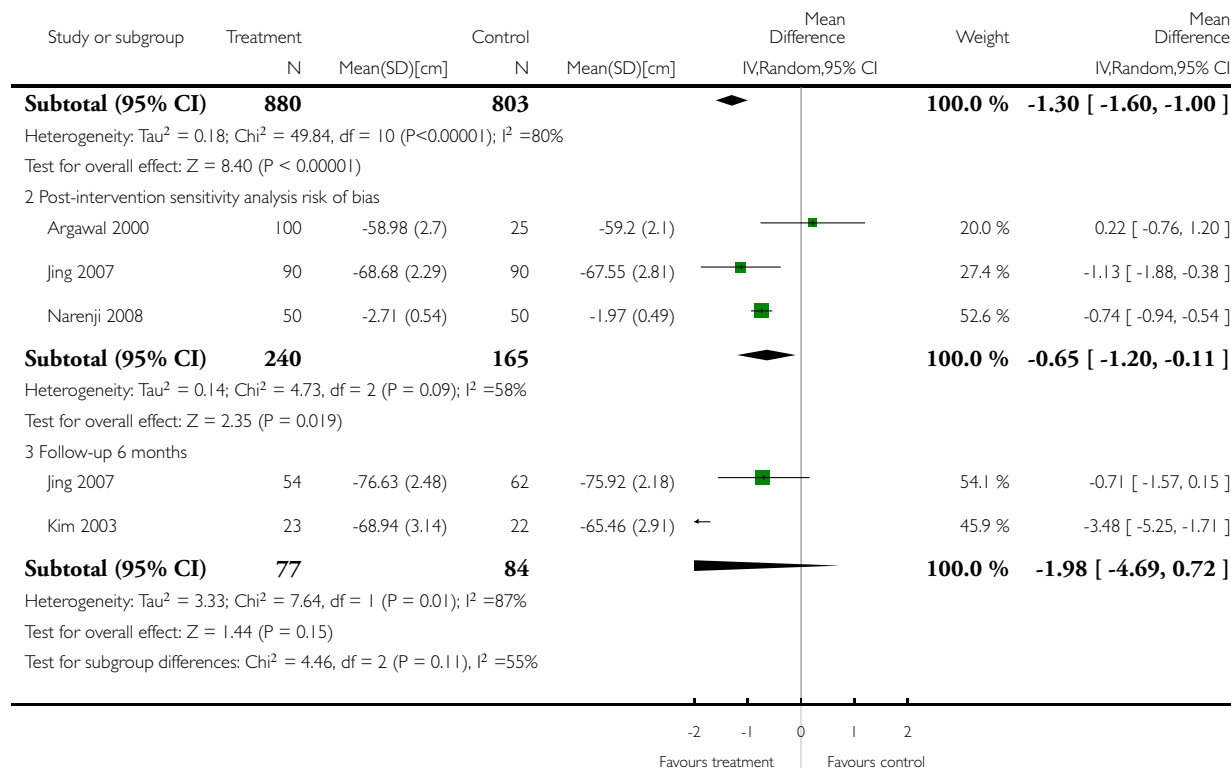
Comparison: 1 Infant massage versus control - physical development

Outcome: 3 Length



(Continued ...)

(... Continued)

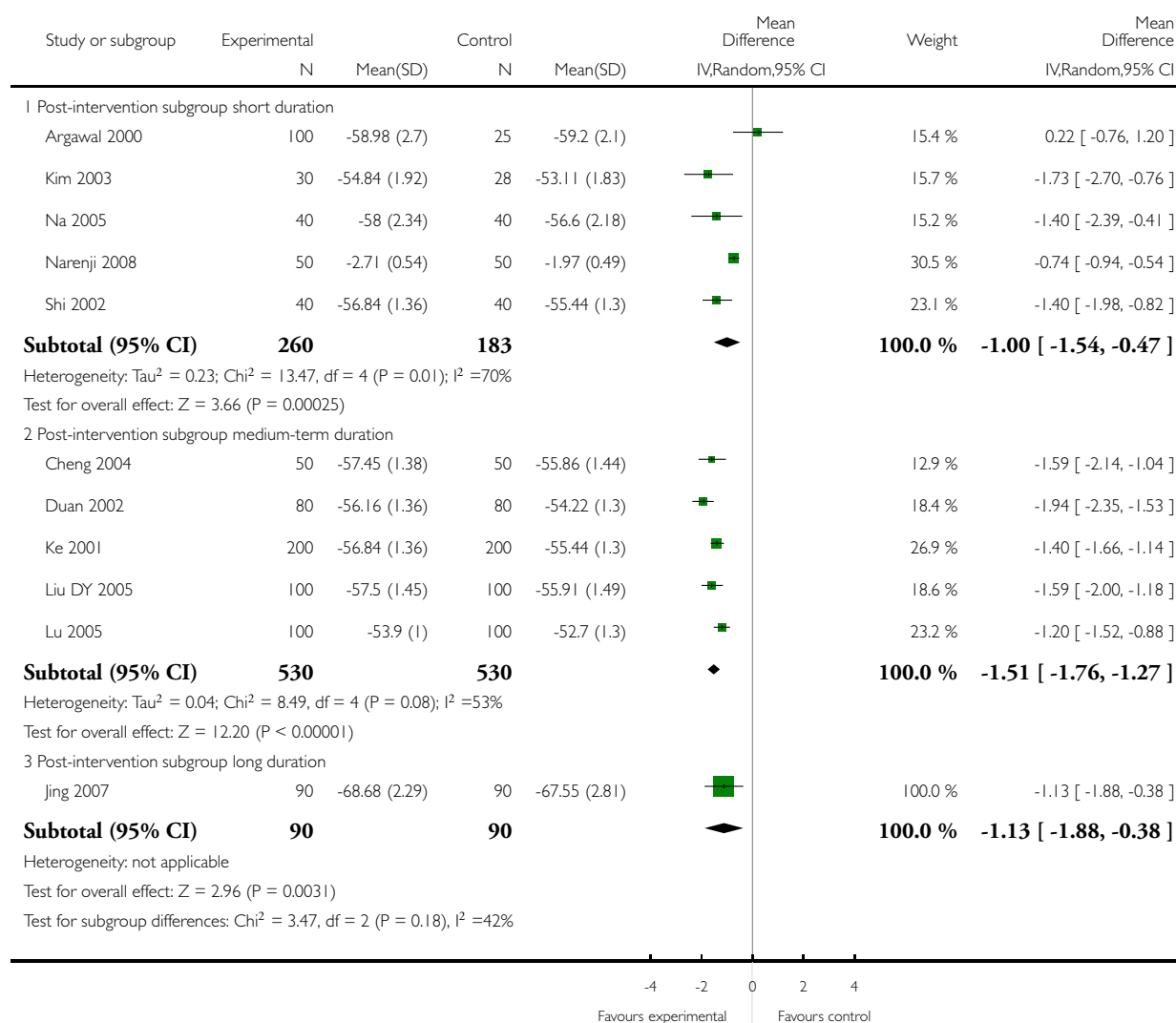


### Analysis 1.4. Comparison 1 Infant massage versus control - physical development, Outcome 4 Length: subgroup analyses (duration of intervention).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 4 Length: subgroup analyses (duration of intervention)

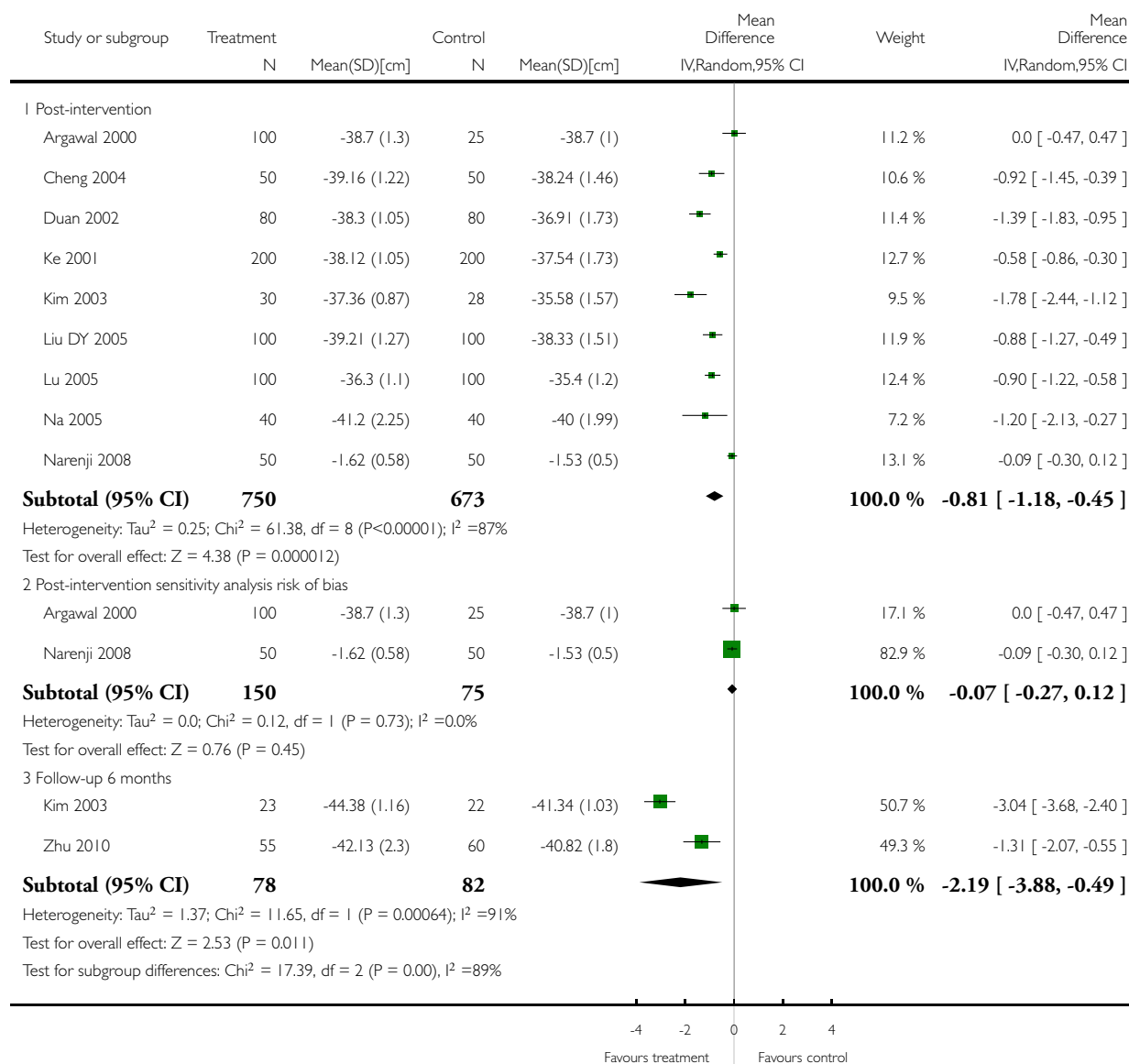


## Analysis 1.5. Comparison 1 Infant massage versus control - physical development, Outcome 5 Head circumference.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 5 Head circumference

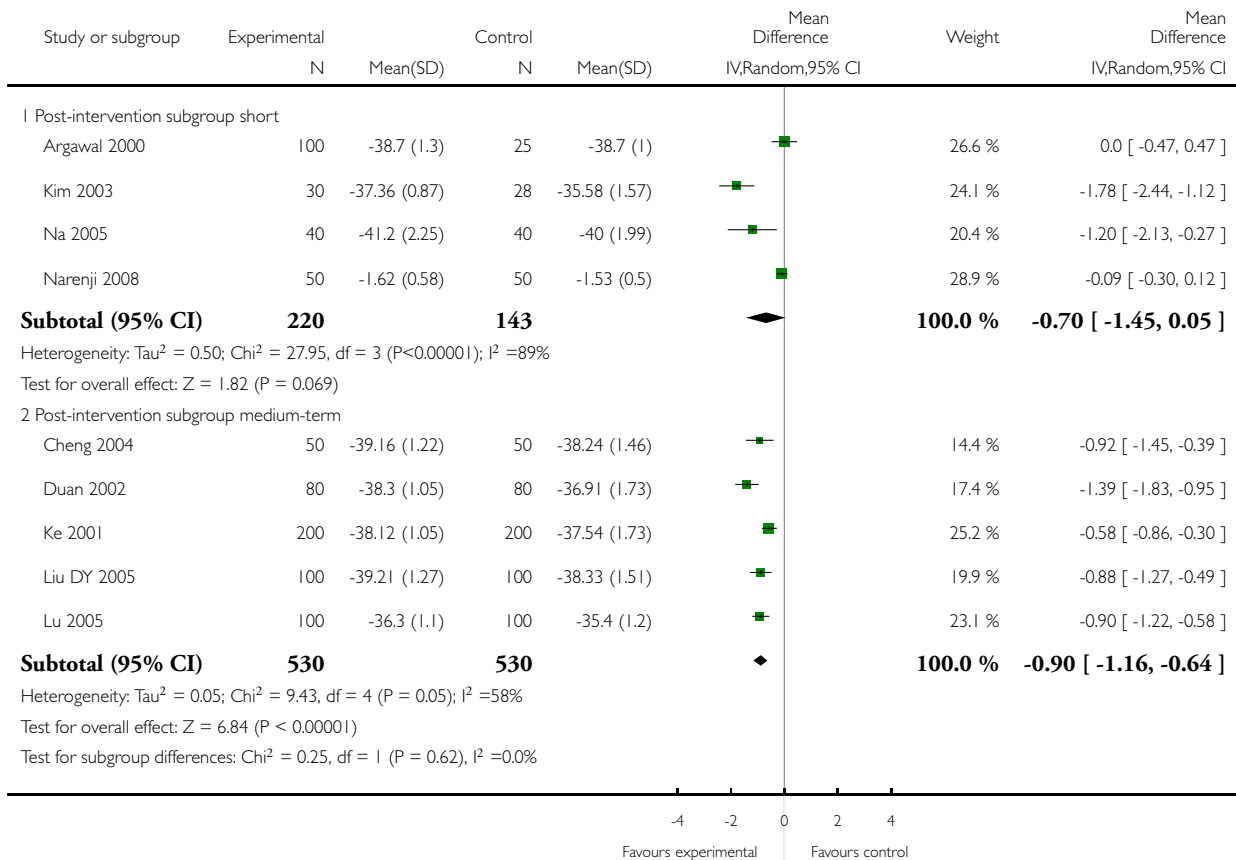


### Analysis 1.6. Comparison 1 Infant massage versus control - physical development, Outcome 6 Head circumference: subgroup analyses (duration of intervention).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 6 Head circumference: subgroup analyses (duration of intervention)

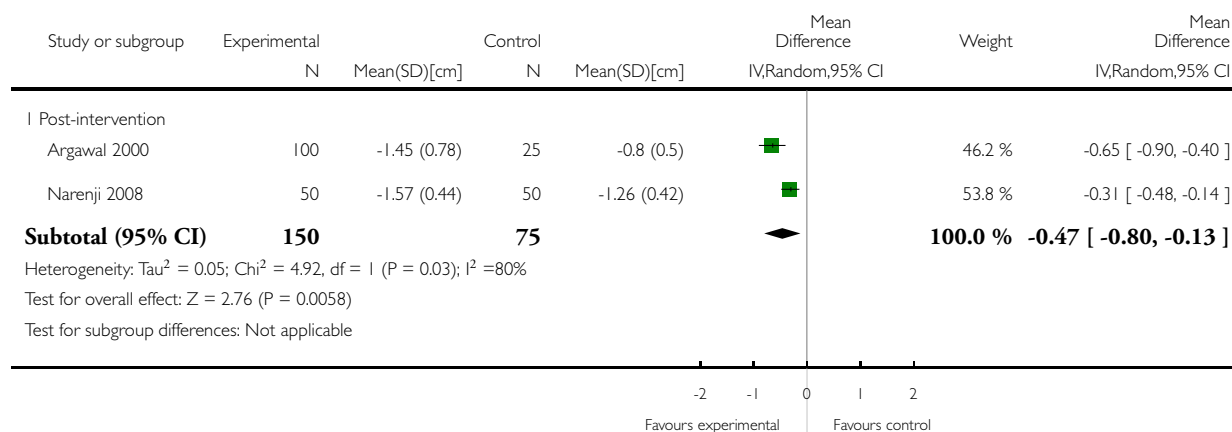


### Analysis 1.7. Comparison 1 Infant massage versus control - physical development, Outcome 7 Mid arm circumference.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 7 Mid arm circumference

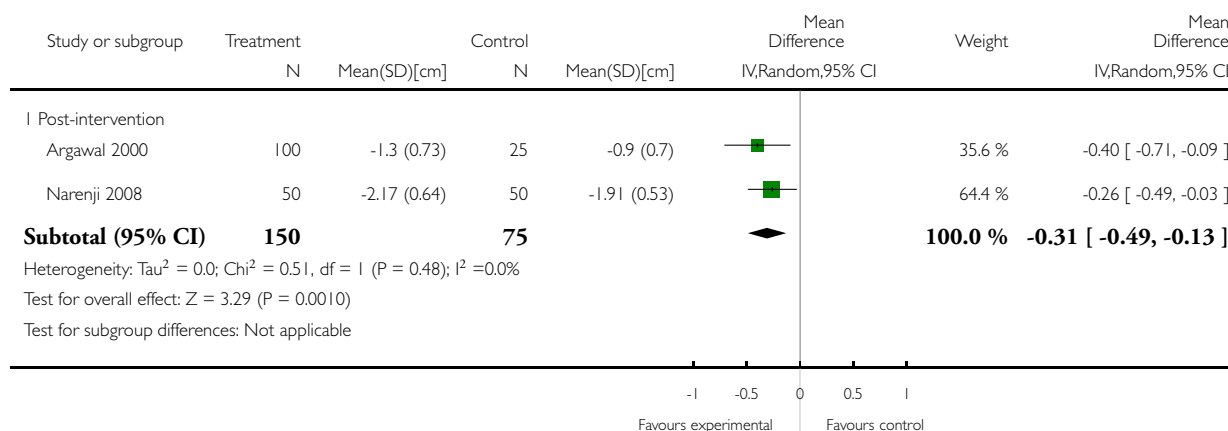


### Analysis 1.8. Comparison 1 Infant massage versus control - physical development, Outcome 8 Mid leg/thigh circumference.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 8 Mid leg/thigh circumference

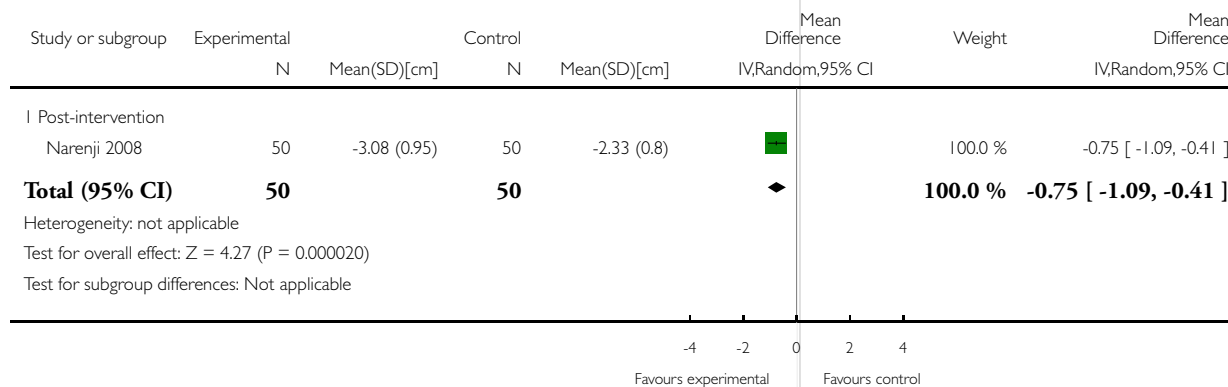


### Analysis 1.9. Comparison 1 Infant massage versus control - physical development, Outcome 9 Abdominal circumference.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 9 Abdominal circumference



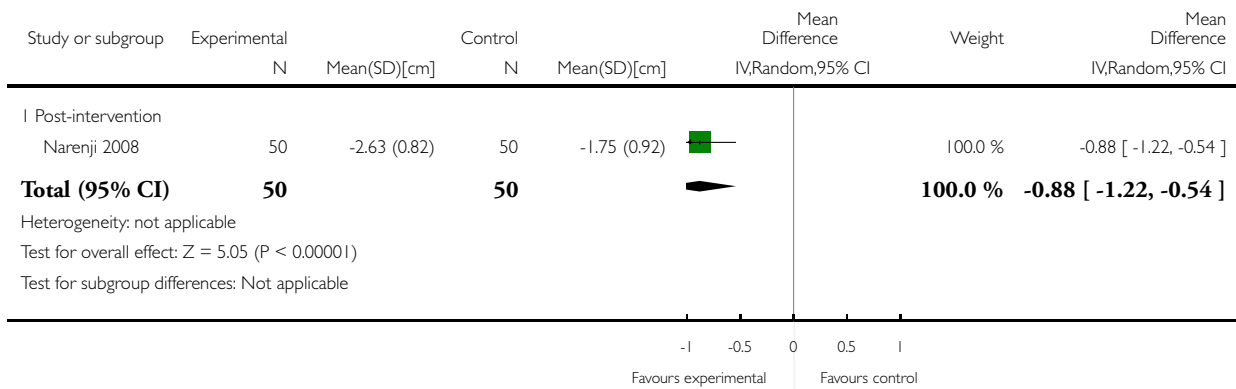


**Analysis 1.10. Comparison 1 Infant massage versus control - physical development, Outcome 10 Chest circumference.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 10 Chest circumference

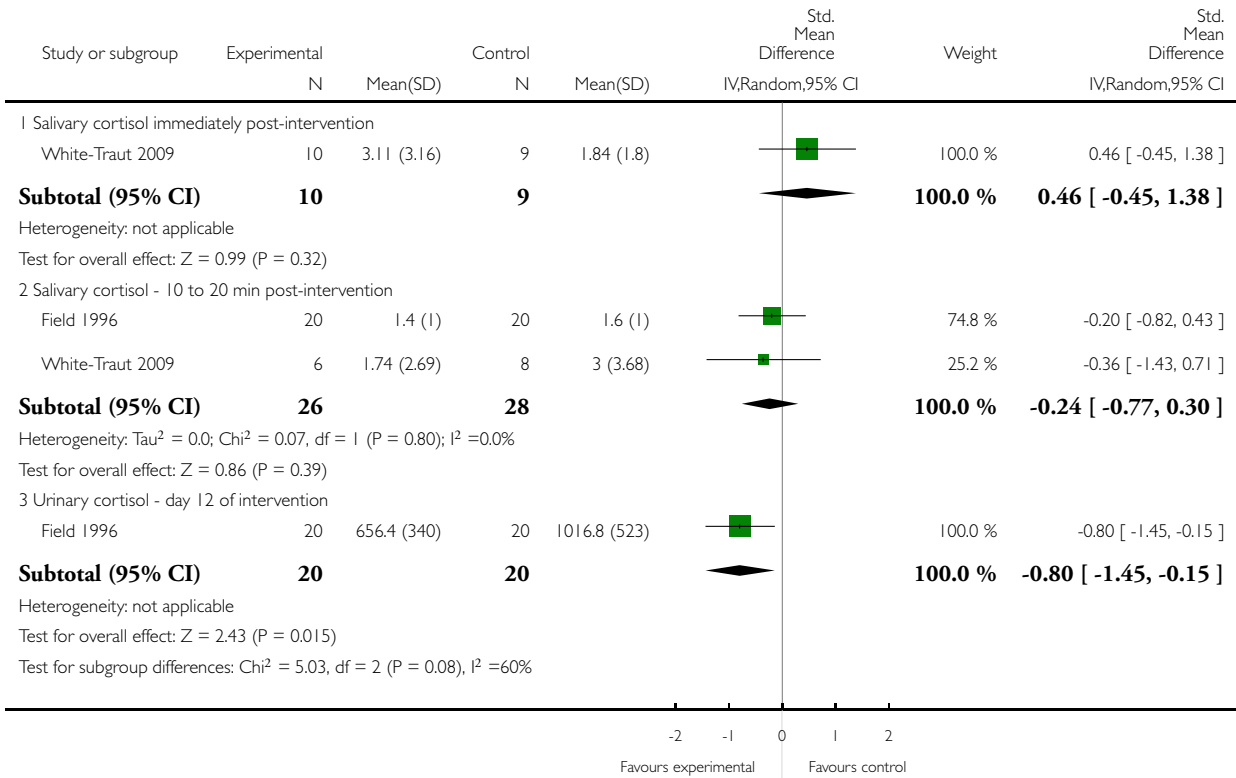


**Analysis 1.11. Comparison 1 Infant massage versus control - physical development, Outcome 11 Hormones: cortisol.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 11 Hormones: cortisol

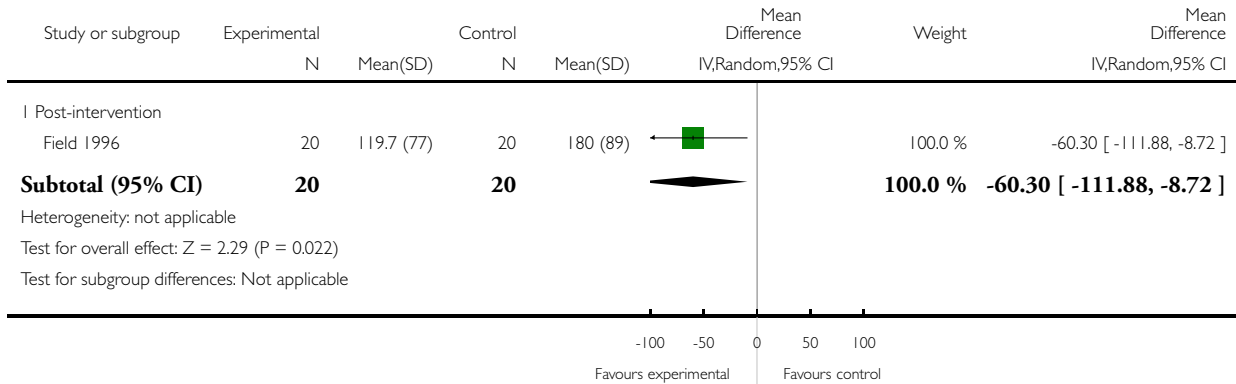


### Analysis 1.12. Comparison 1 Infant massage versus control - physical development, Outcome 12 Hormones: norepinephrine.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 12 Hormones: norepinephrine

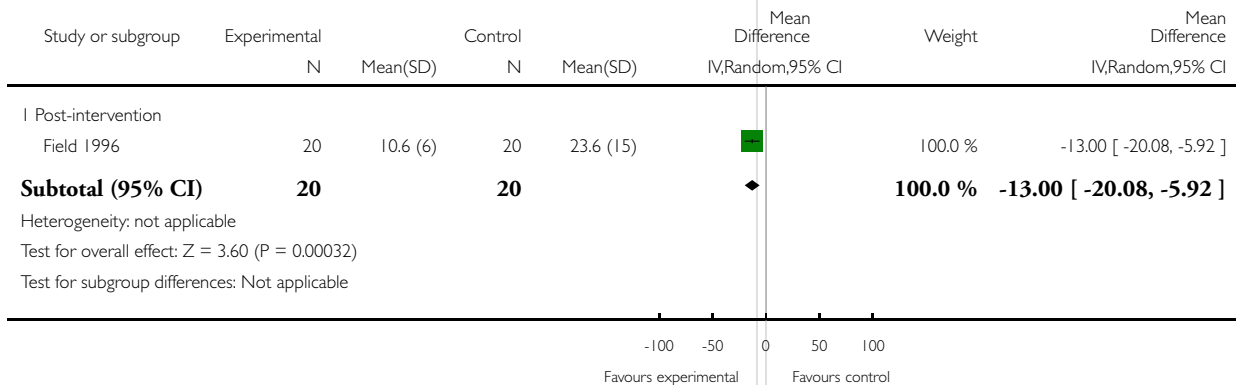


### Analysis 1.13. Comparison 1 Infant massage versus control - physical development, Outcome 13 Hormones: epinephrine.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 13 Hormones: epinephrine

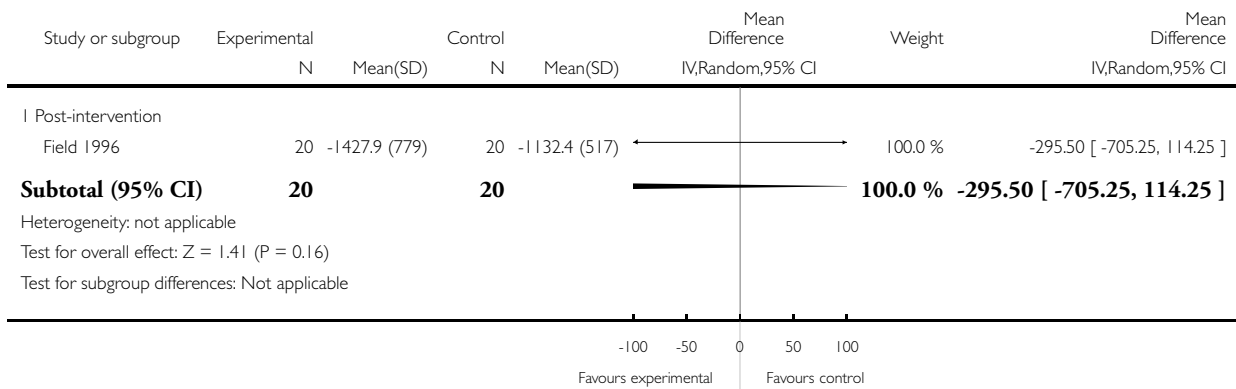


**Analysis 1.14. Comparison 1 Infant massage versus control - physical development, Outcome 14 Hormones: serotonin.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 14 Hormones: serotonin

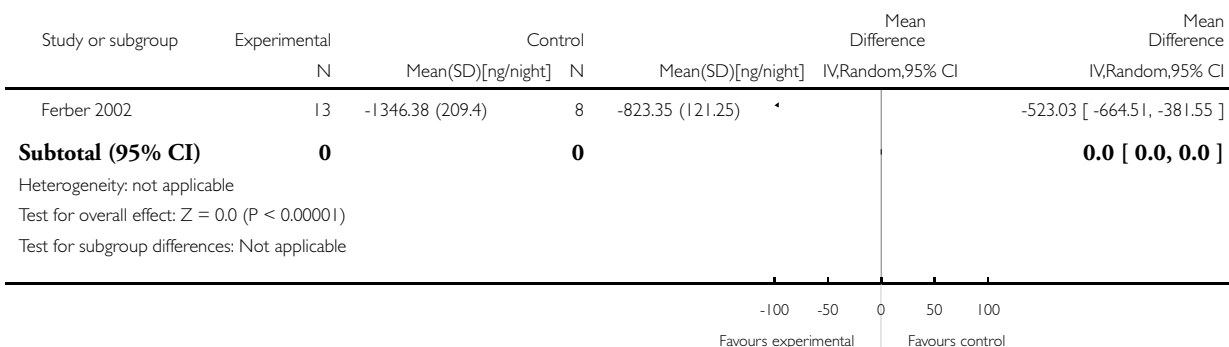


**Analysis 1.15. Comparison 1 Infant massage versus control - physical development, Outcome 15 Hormones: 6-sulphatoxymelatonin secretion.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 15 Hormones: 6-sulphatoxymelatonin secretion

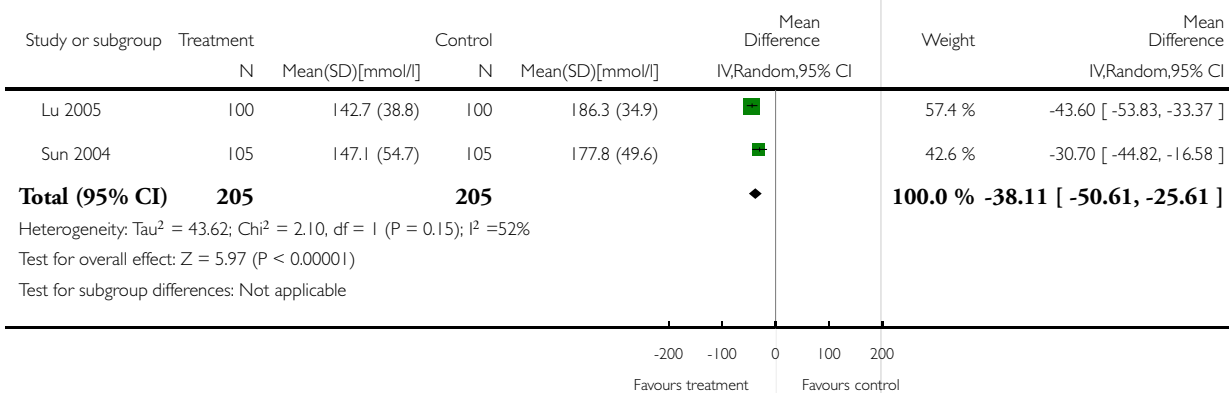


**Analysis 1.16. Comparison 1 Infant massage versus control - physical development, Outcome 16 Biochemical markers: Bilirubin (7 days PN).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 16 Biochemical markers: Bilirubin (7 days PN)

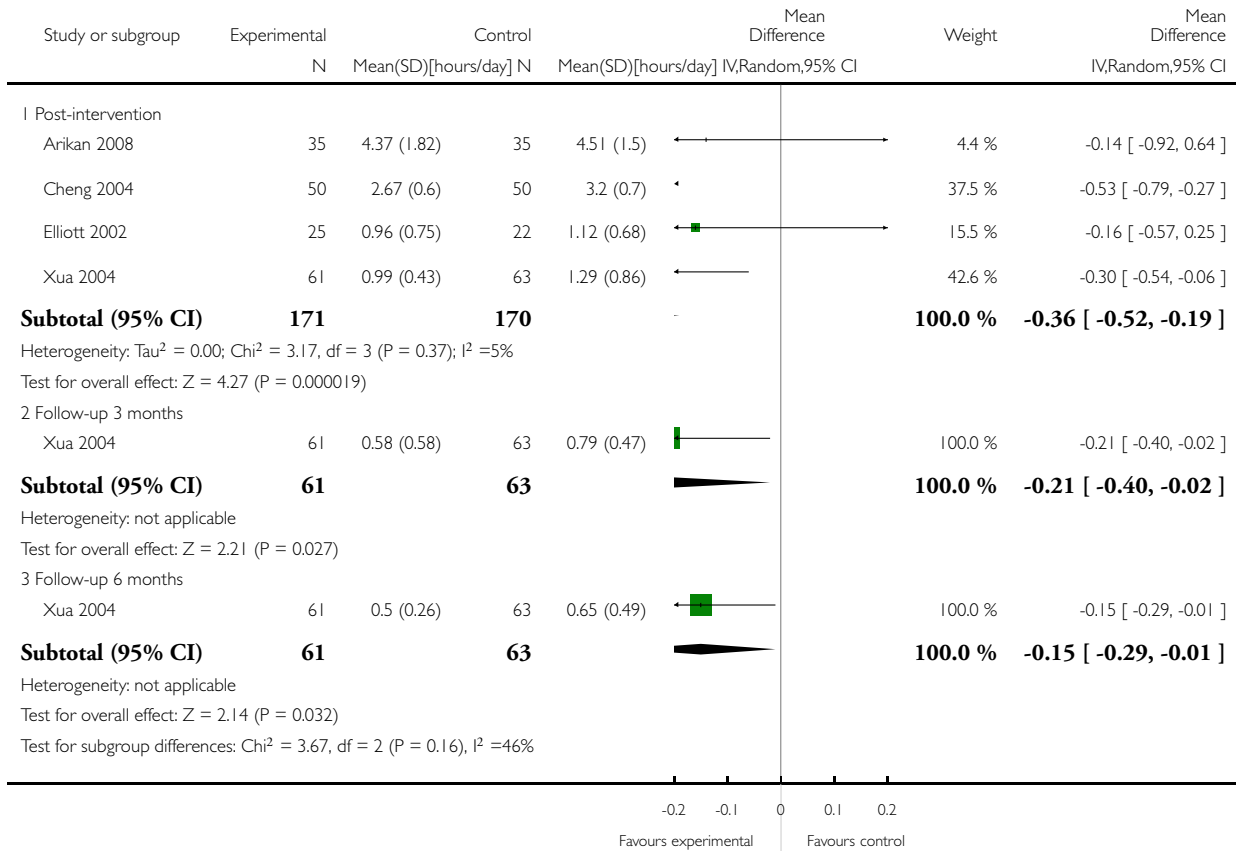


### Analysis 1.17. Comparison 1 Infant massage versus control - physical development, Outcome 17 Crying or fussing time.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 17 Crying or fussing time

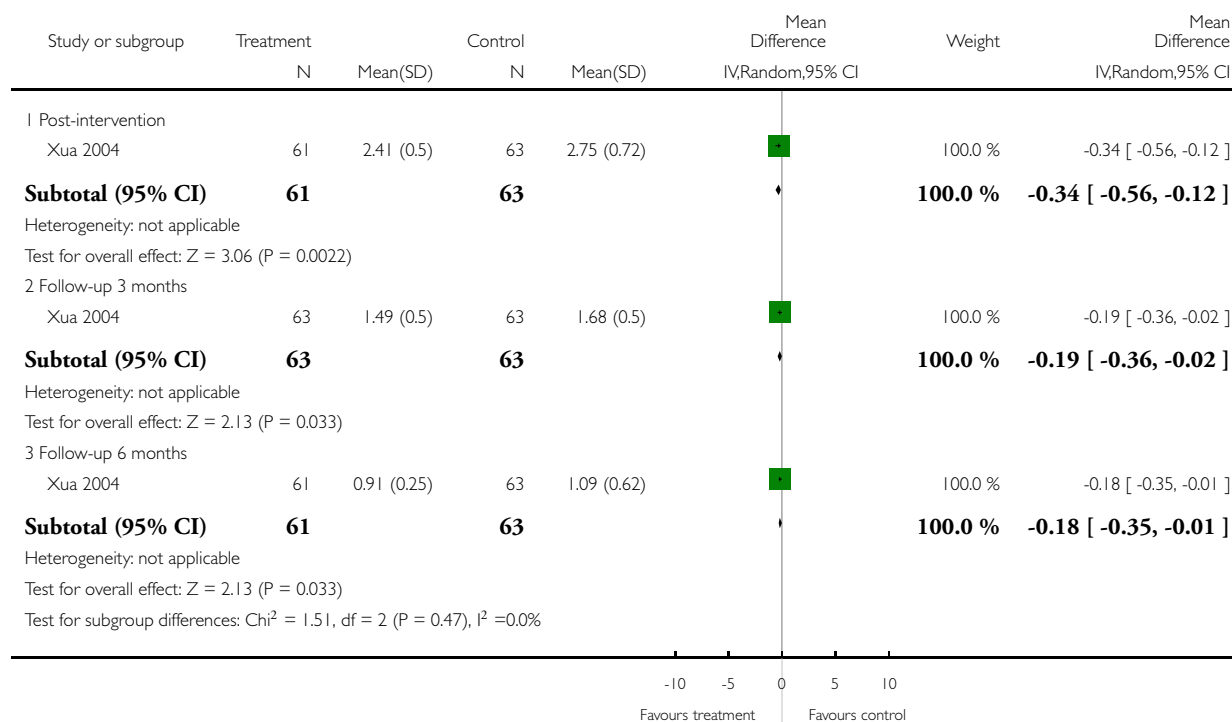


### Analysis 1.18. Comparison 1 Infant massage versus control - physical development, Outcome 18 Crying frequency (times).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 18 Crying frequency (times)

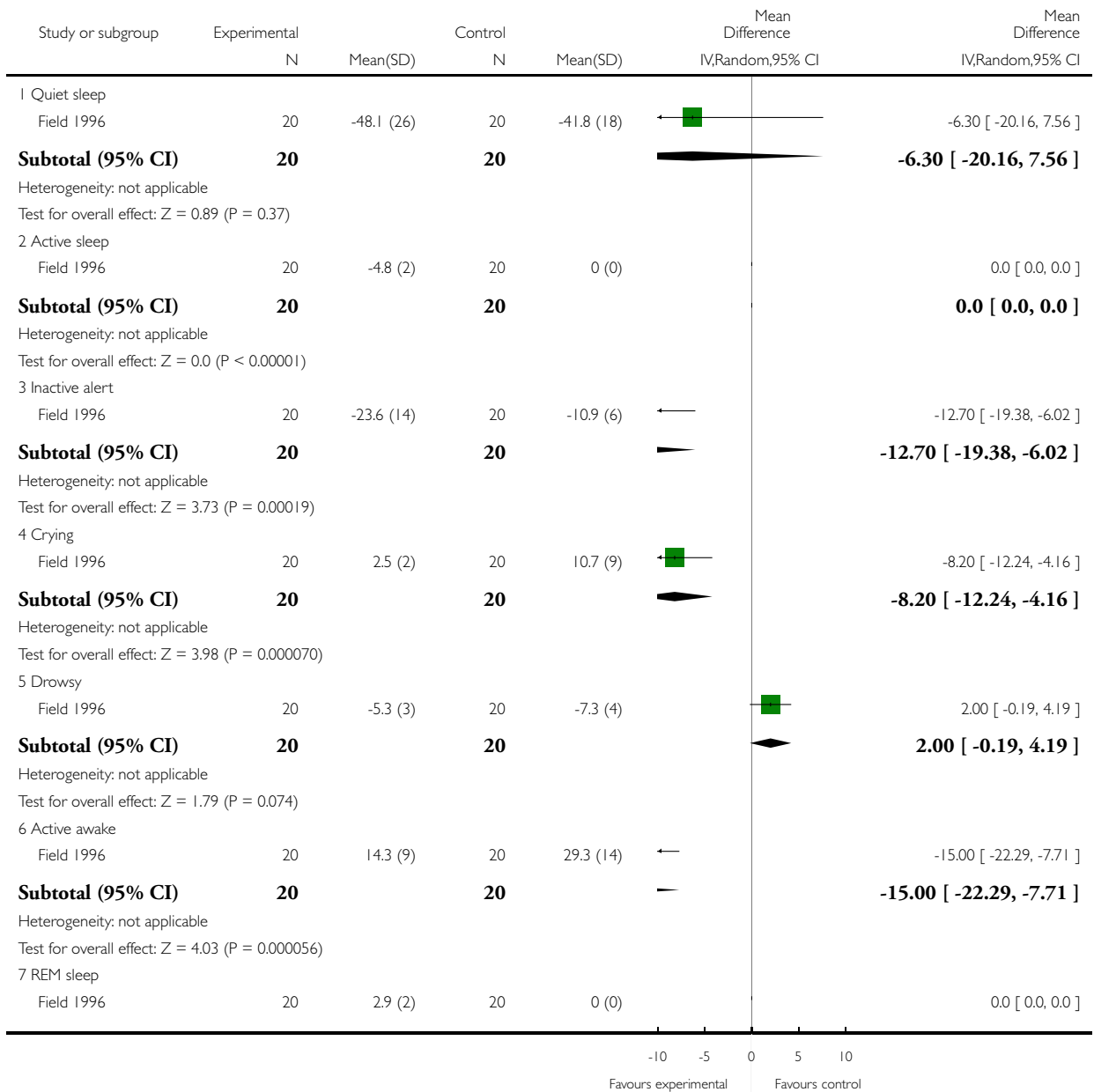


### Analysis 1.19. Comparison 1 Infant massage versus control - physical development, Outcome 19 Sleep/wake behaviours (Thoman).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

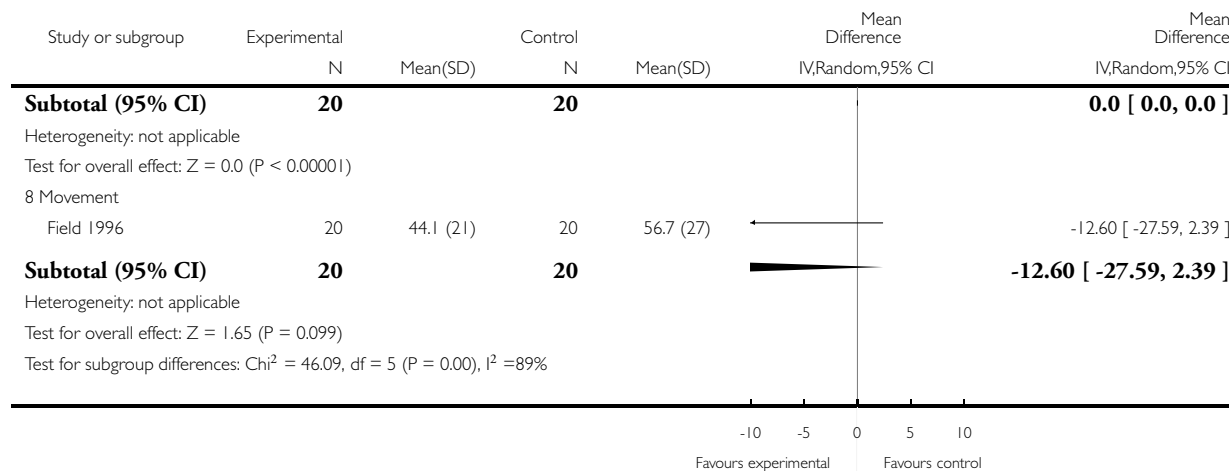
Outcome: 19 Sleep/wake behaviours (Thoman)



(Continued ...)



(... Continued)

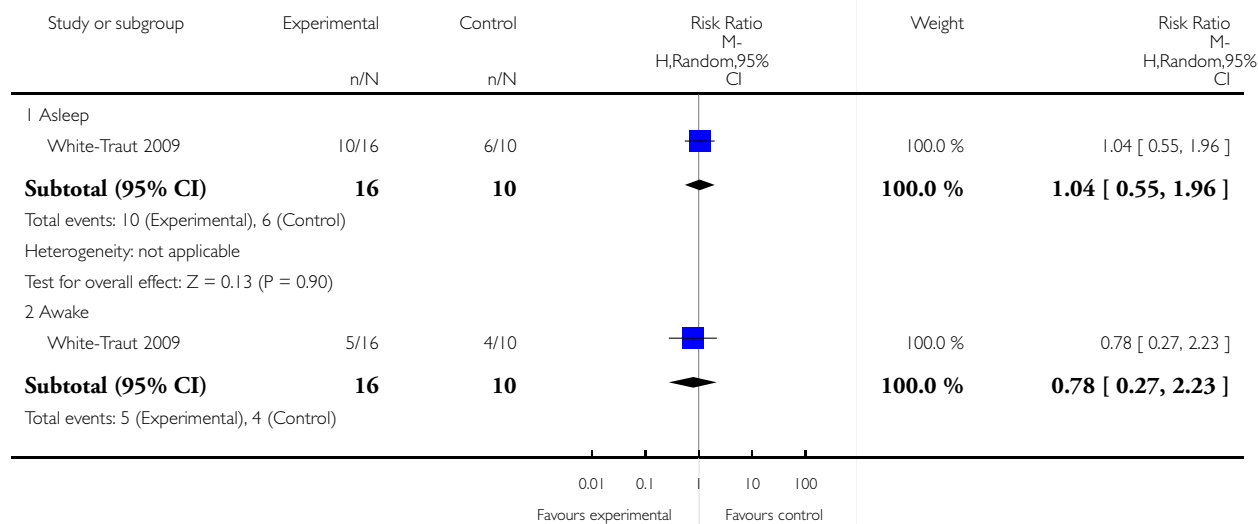


### Analysis 1.20. Comparison 1 Infant massage versus control - physical development, Outcome 20 Behavioural state immediately post-intervention (Thoman).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

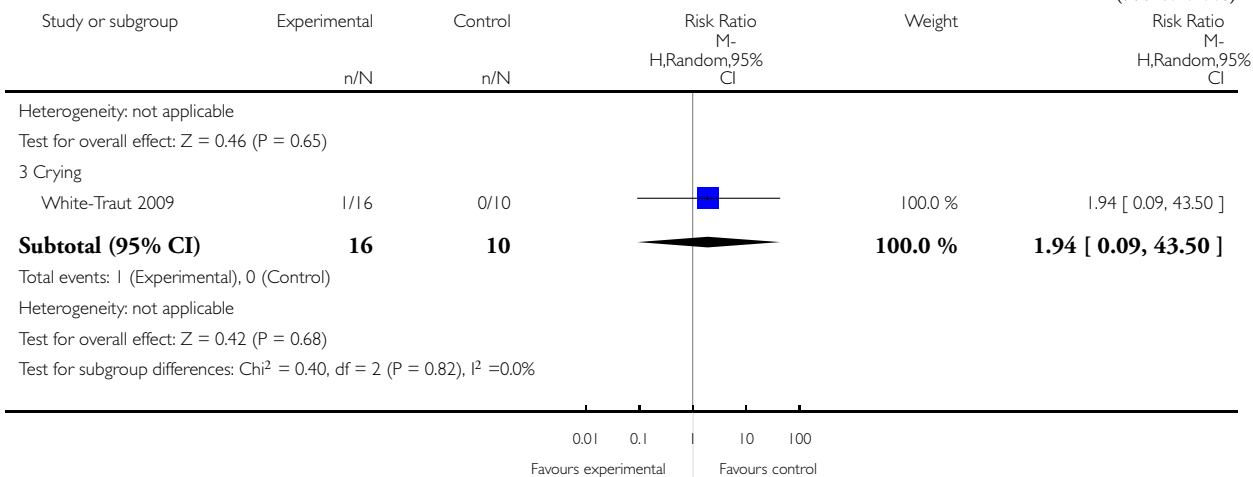
Comparison: 1 Infant massage versus control - physical development

Outcome: 20 Behavioural state immediately post-intervention (Thoman)



(Continued ...)

(... Continued)

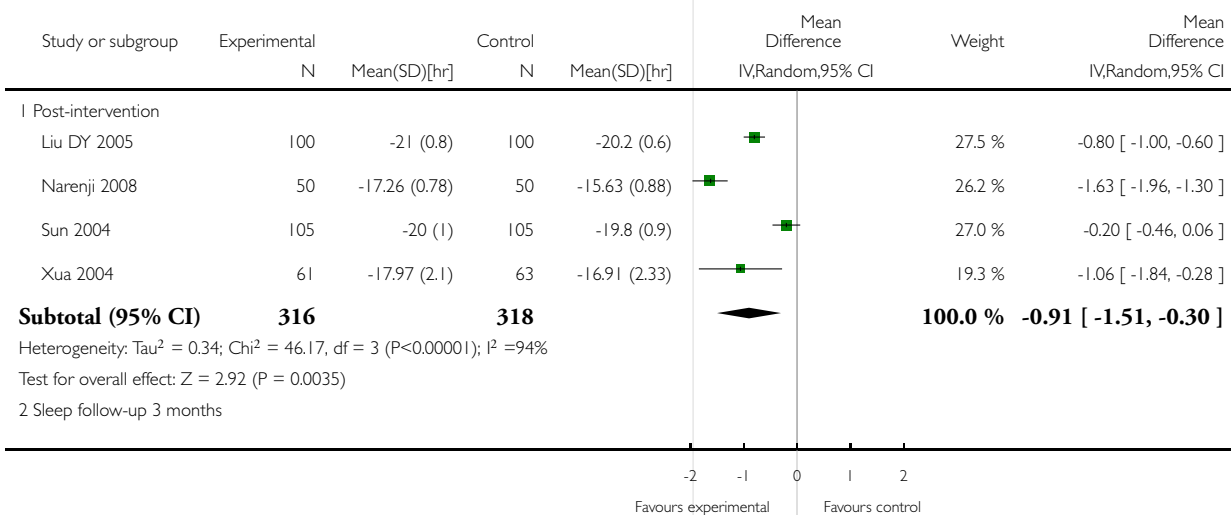


### Analysis 1.21. Comparison 1 Infant massage versus control - physical development, Outcome 21 Sleep duration over 24hr period.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

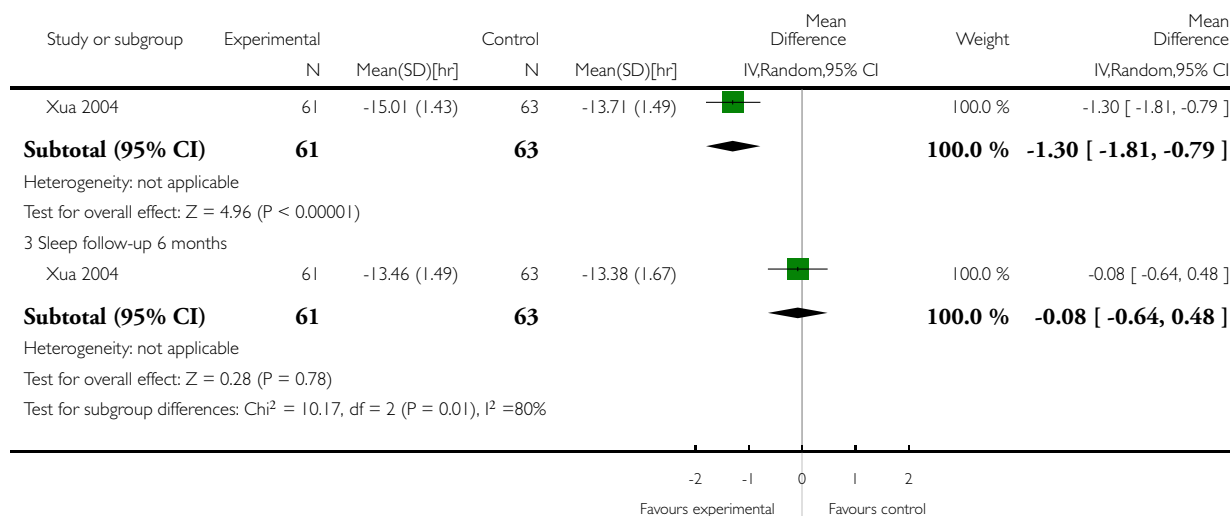
Comparison: 1 Infant massage versus control - physical development

Outcome: 21 Sleep duration over 24hr period



(Continued ...)

(... Continued)

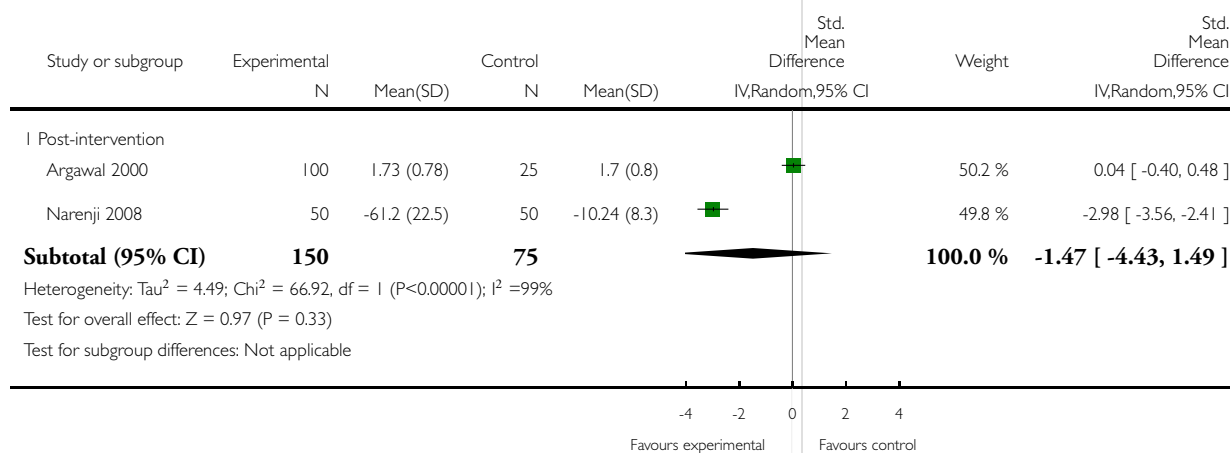


### Analysis 1.22. Comparison 1 Infant massage versus control - physical development, Outcome 22 Mean increase in 24h sleep.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 22 Mean increase in 24h sleep

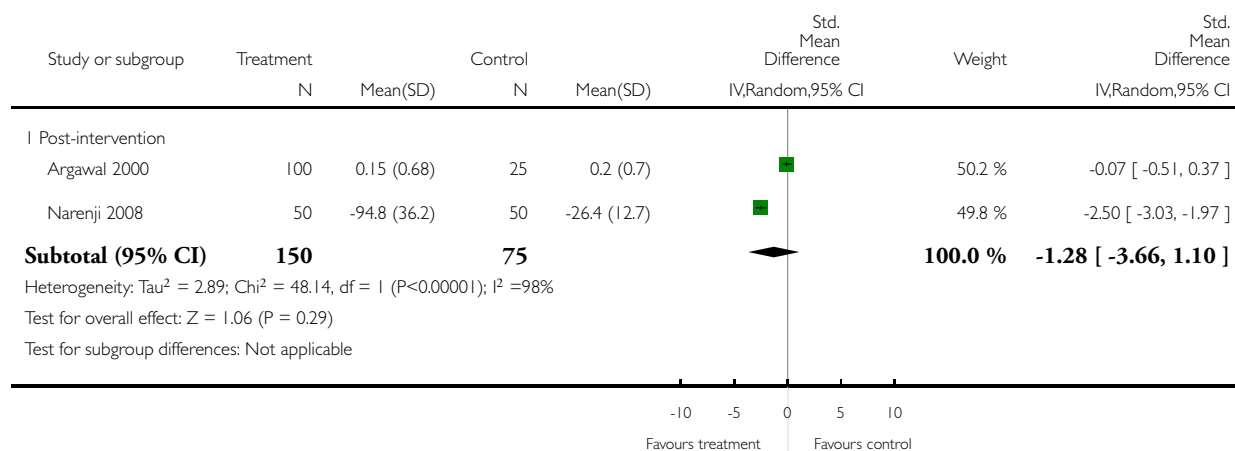


### Analysis 1.23. Comparison I Infant massage versus control - physical development, Outcome 23 Mean increase in duration of night sleep.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: I Infant massage versus control - physical development

Outcome: 23 Mean increase in duration of night sleep

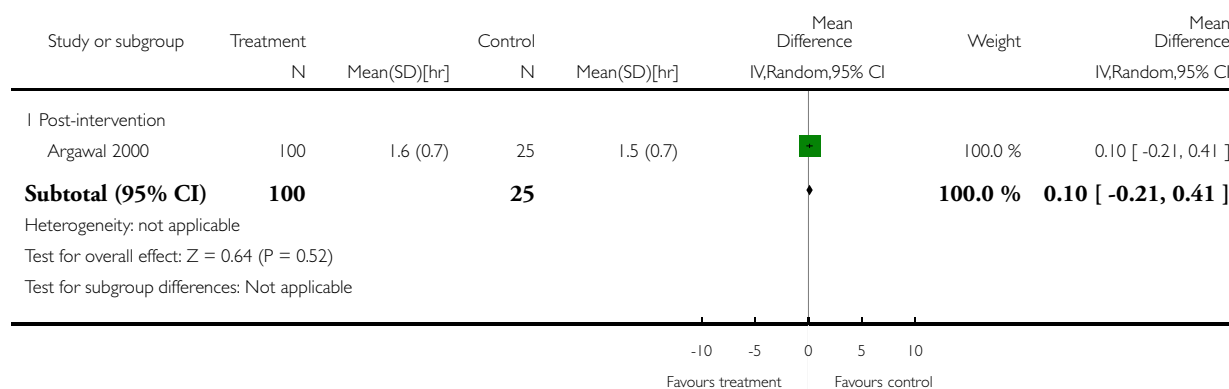


### Analysis 1.24. Comparison I Infant massage versus control - physical development, Outcome 24 Mean increase in duration of day sleep.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: I Infant massage versus control - physical development

Outcome: 24 Mean increase in duration of day sleep

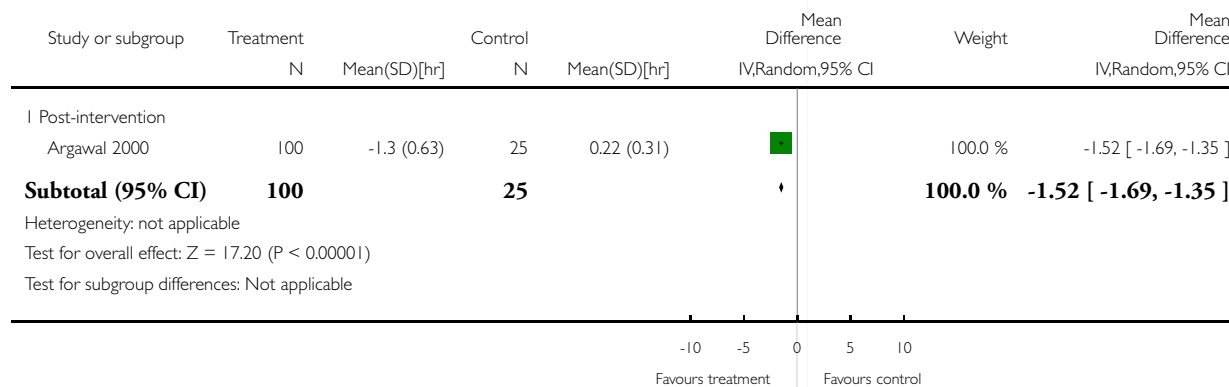


### Analysis 1.25. Comparison I Infant massage versus control - physical development, Outcome 25 Mean increase in duration of first morning sleep after massage.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: I Infant massage versus control - physical development

Outcome: 25 Mean increase in duration of first morning sleep after massage

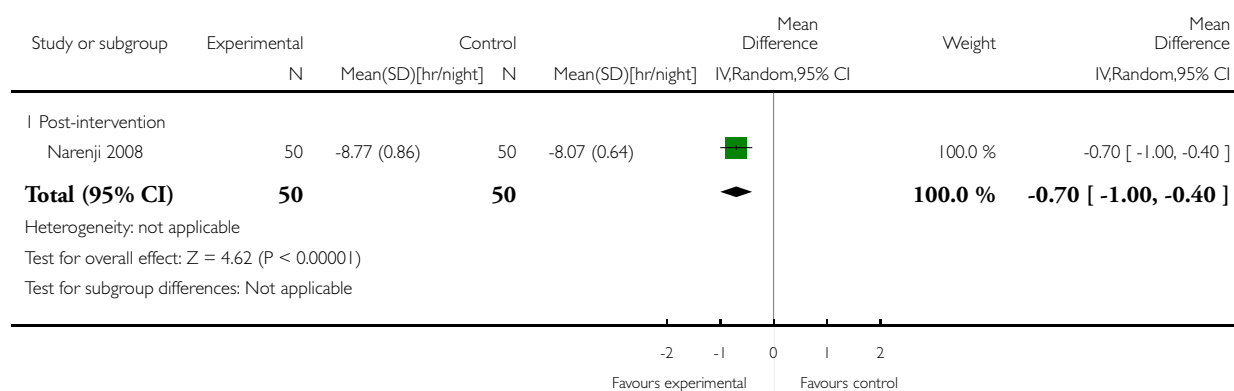


### Analysis 1.26. Comparison 1 Infant massage versus control - physical development, Outcome 26 Sleep (total hours per night).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 26 Sleep (total hours per night)

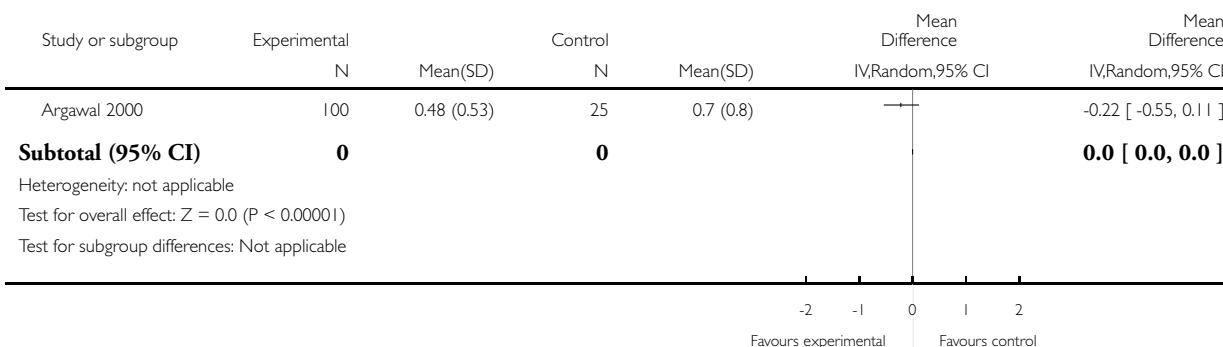


### Analysis 1.27. Comparison 1 Infant massage versus control - physical development, Outcome 27 Number of naps (total number of naps).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 27 Number of naps (total number of naps)

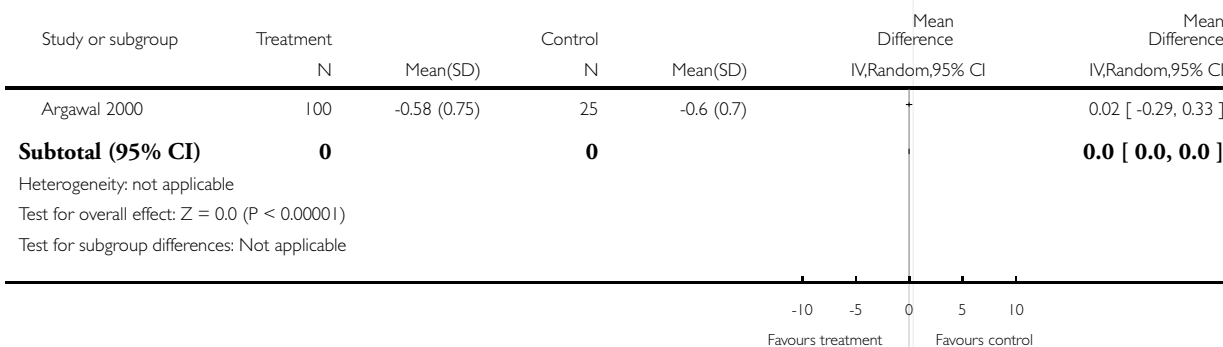


### Analysis 1.28. Comparison 1 Infant massage versus control - physical development, Outcome 28 Number of naps in day.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 28 Number of naps in day

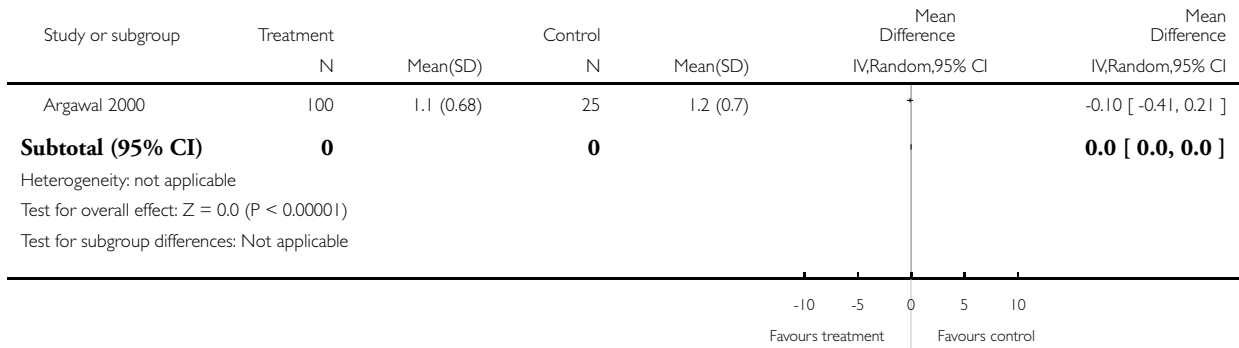


**Analysis 1.29. Comparison 1 Infant massage versus control - physical development, Outcome 29 Number of naps at night.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 29 Number of naps at night



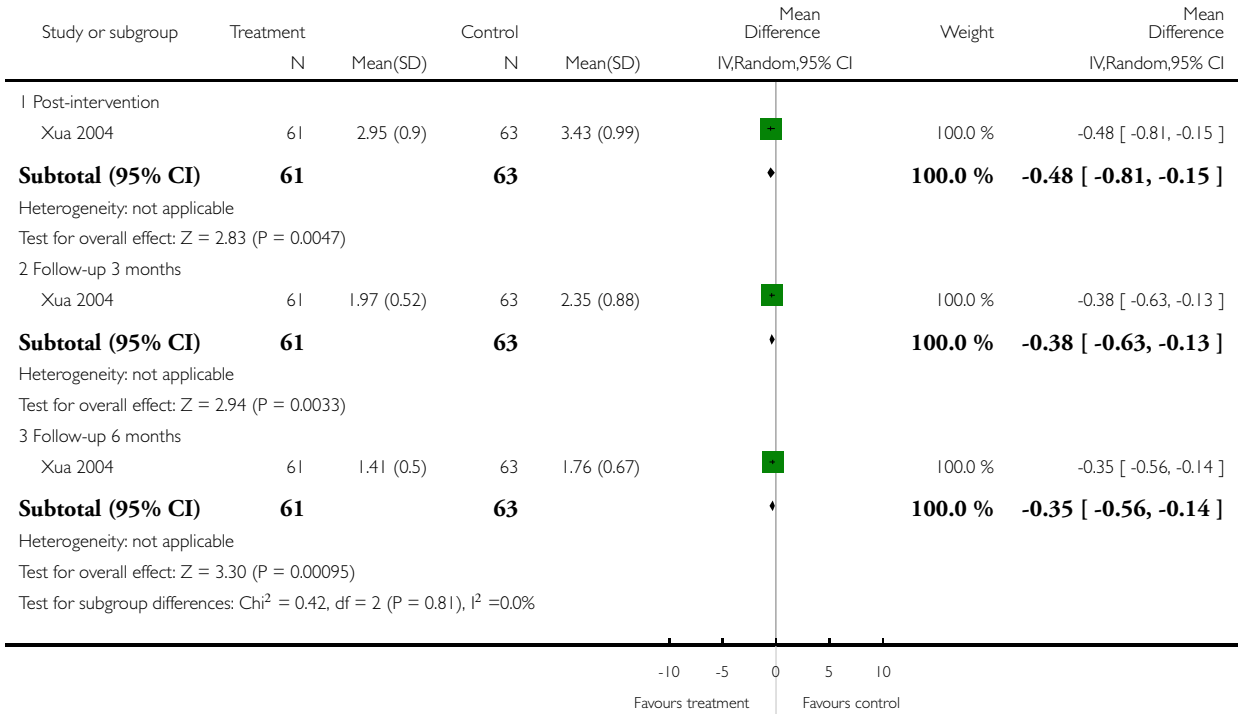


### Analysis 1.30. Comparison 1 Infant massage versus control - physical development, Outcome 30 Night Wake Frequency (times).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 30 Night Wake Frequency (times)

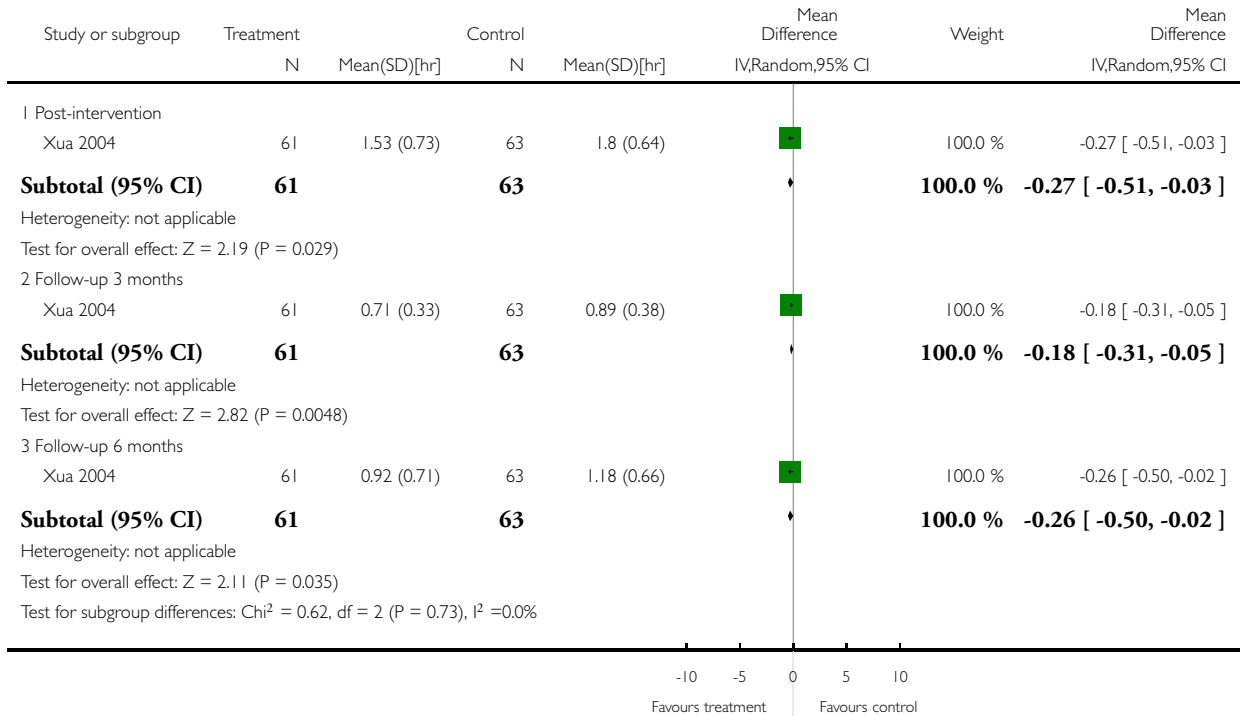


### Analysis 1.31. Comparison 1 Infant massage versus control - physical development, Outcome 31 Night wake duration.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 31 Night wake duration

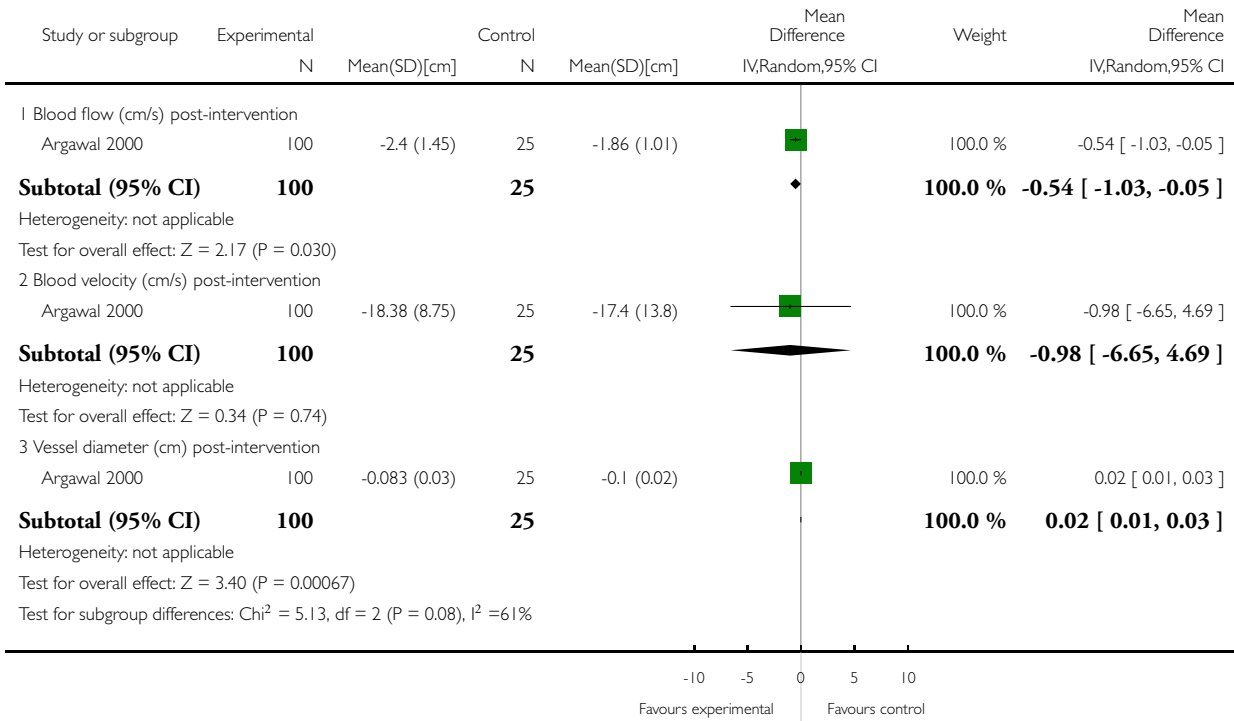


**Analysis 1.32. Comparison 1 Infant massage versus control - physical development, Outcome 32 Blood flow (post intervention).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 32 Blood flow (post intervention)

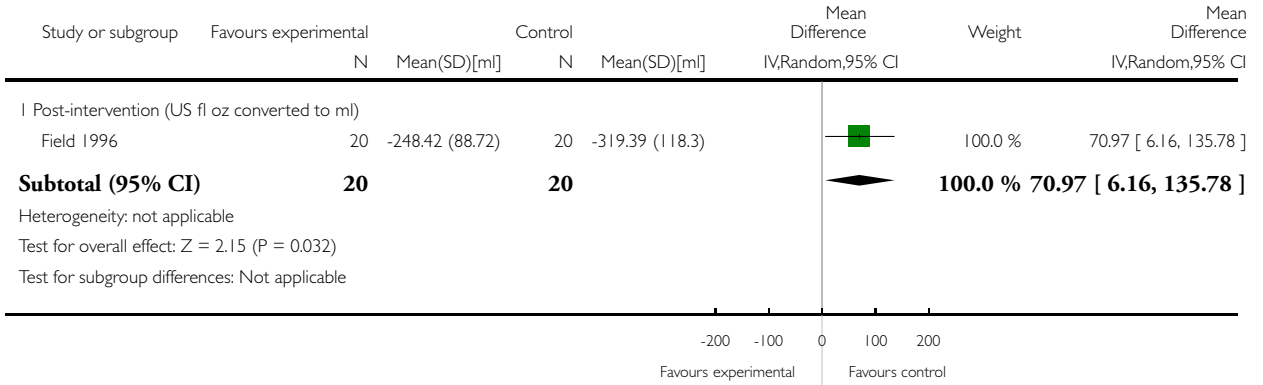


### Analysis 1.33. Comparison 1 Infant massage versus control - physical development, Outcome 33 Formula intake.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 33 Formula intake

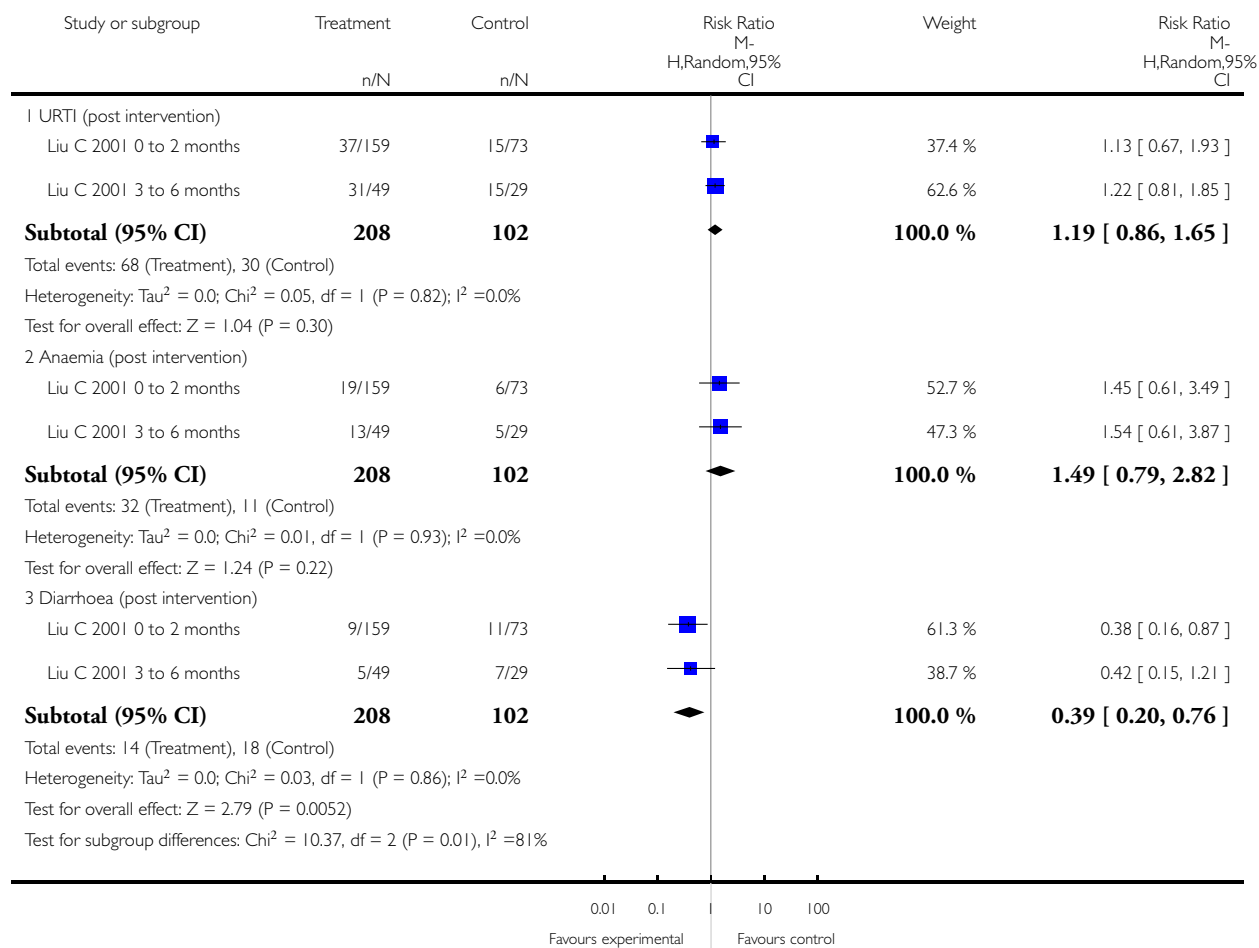


### Analysis 1.34. Comparison 1 Infant massage versus control - physical development, Outcome 34 Illness.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 34 Illness

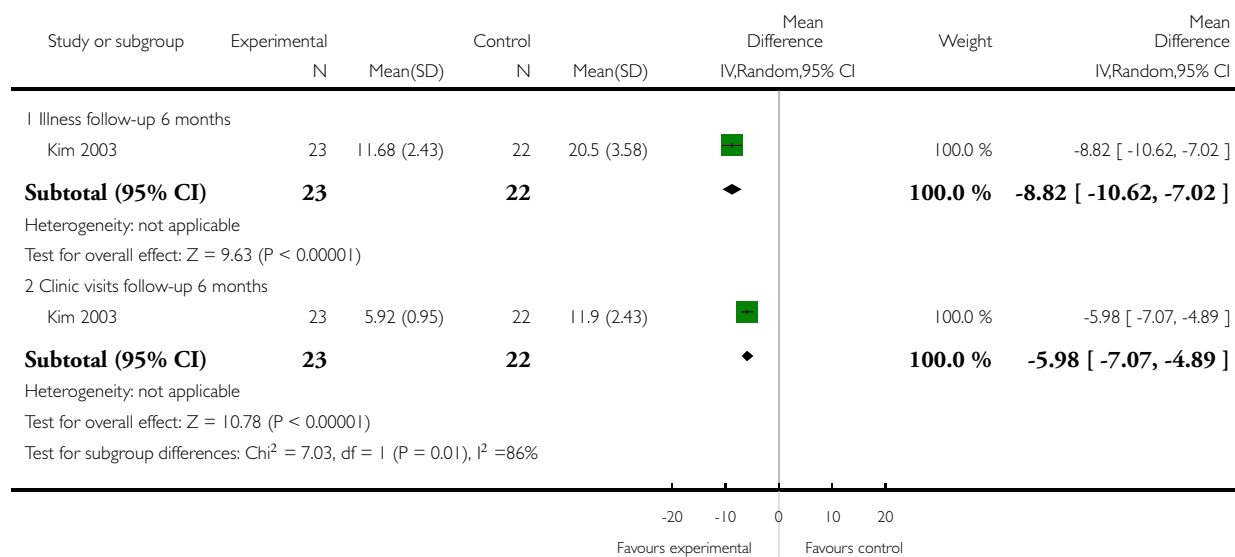


### Analysis 1.35. Comparison 1 Infant massage versus control - physical development, Outcome 35 Illness and clinic visits.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 1 Infant massage versus control - physical development

Outcome: 35 Illness and clinic visits

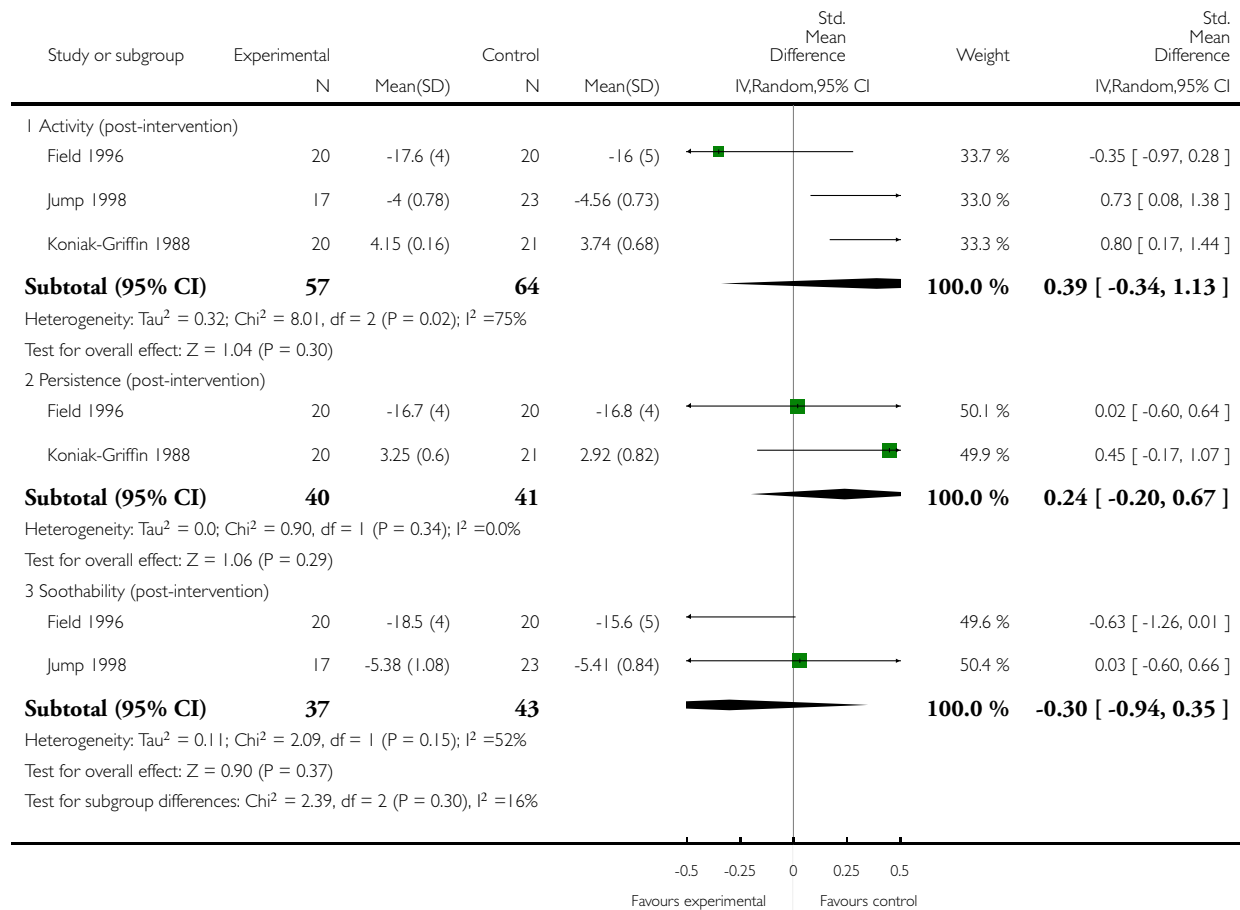


## Analysis 2.1. Comparison 2 Infant massage versus control - mental health and development, Outcome 1 Infant temperament meta-analyses.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 1 Infant temperament meta-analyses

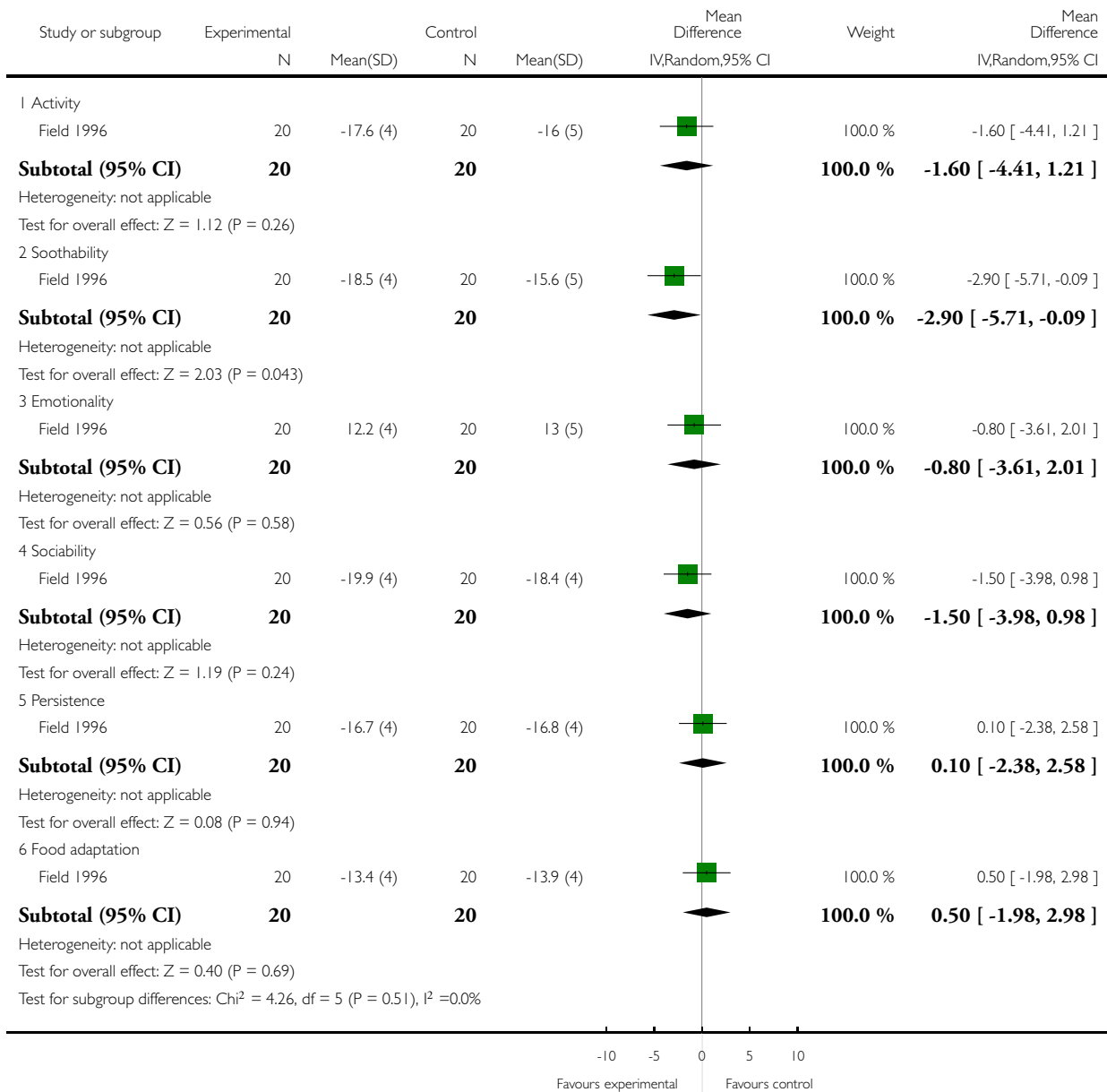


## Analysis 2.2. Comparison 2 Infant massage versus control - mental health and development, Outcome 2 Infant temperament (CCTI) post intervention.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 2 Infant temperament (CCTI) post intervention



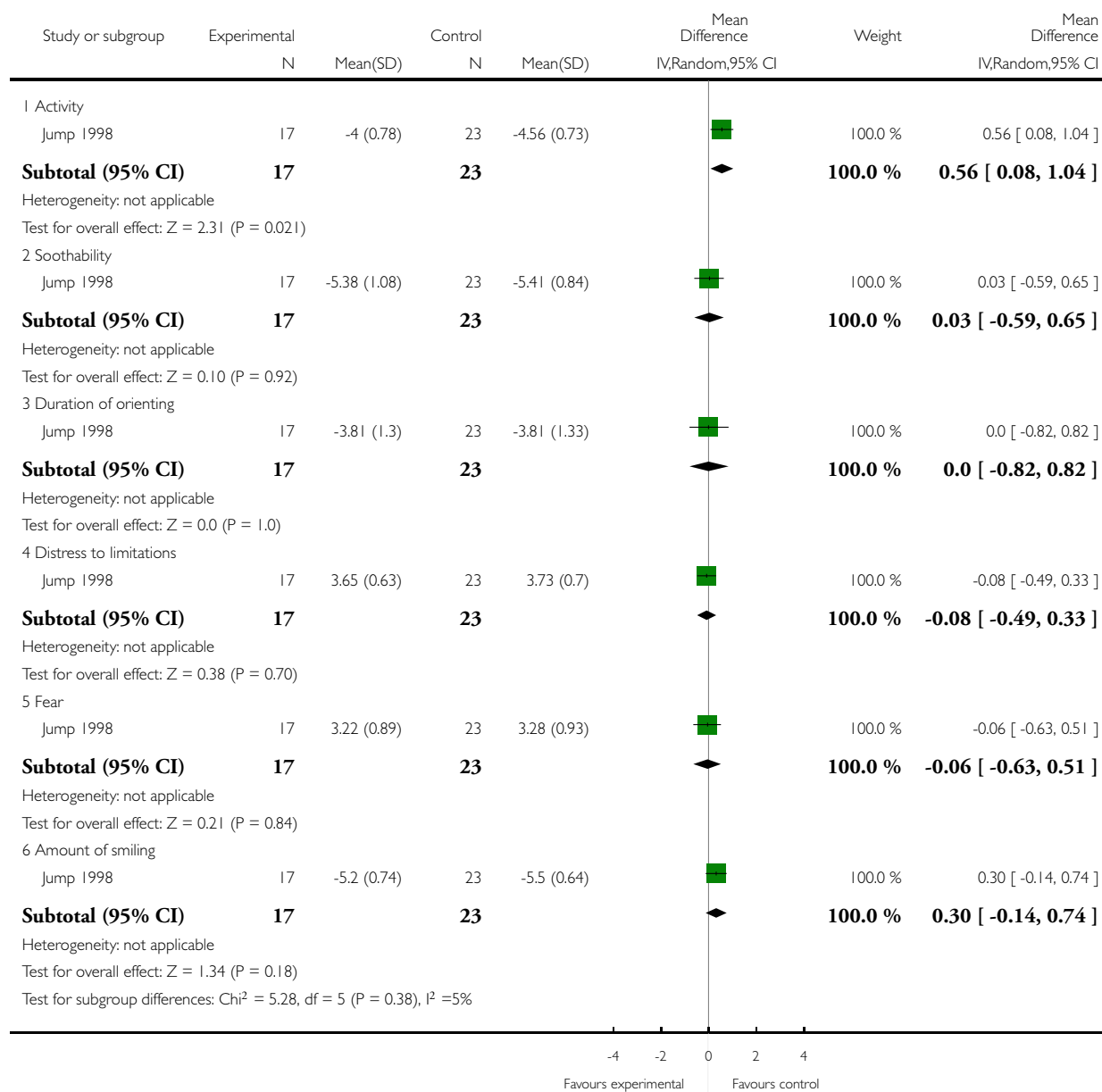


### Analysis 2.3. Comparison 2 Infant massage versus control - mental health and development, Outcome 3 Infant temperament (Infant behaviour questionnaire (IBQ) post intervention).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 3 Infant temperament (Infant behaviour questionnaire (IBQ) post intervention)

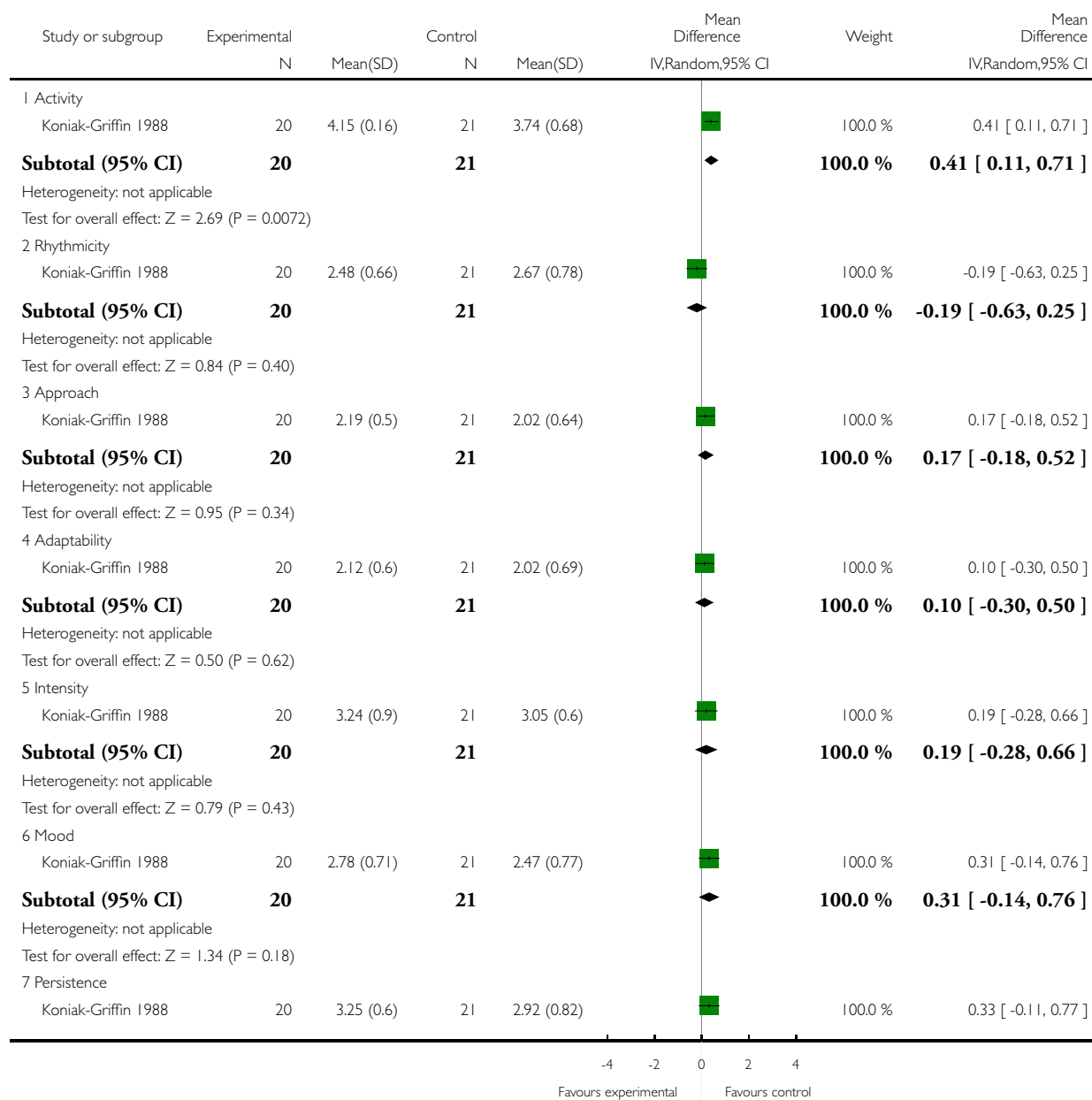


### Analysis 2.4. Comparison 2 Infant massage versus control - mental health and development, Outcome 4 Infant temperament questionnaire (revised RITQ (Carey)) post-intervention 4 months.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

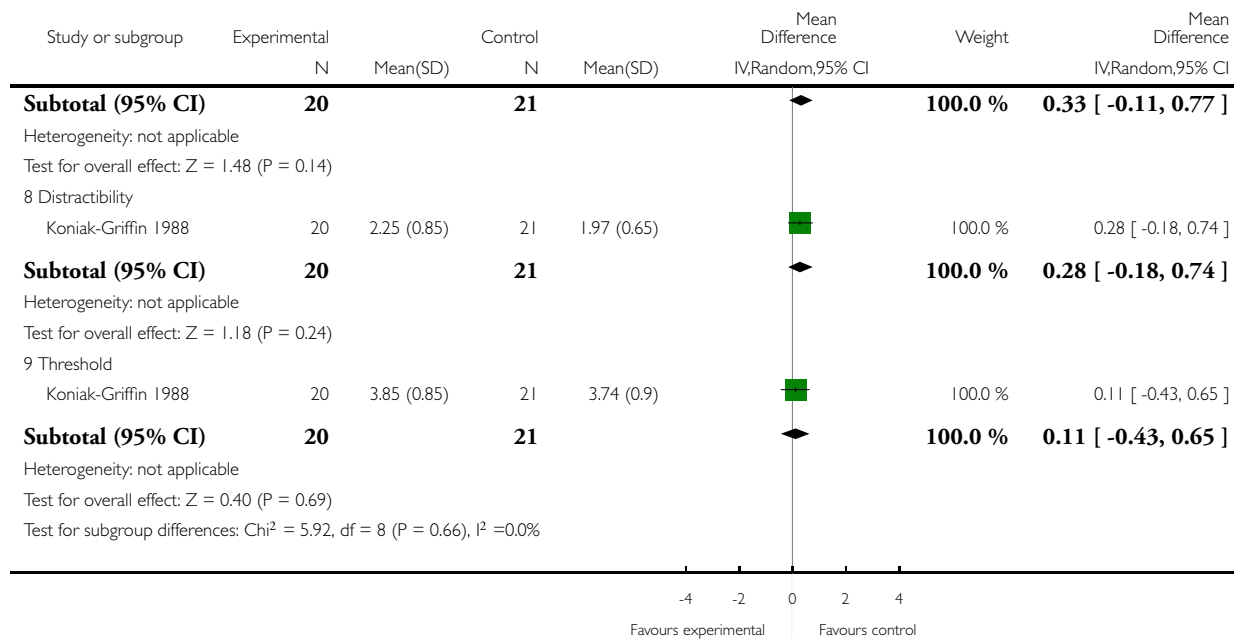
Comparison: 2 Infant massage versus control - mental health and development

Outcome: 4 Infant temperament questionnaire (revised RITQ (Carey)) post-intervention 4 months



(Continued ...)

(... Continued)

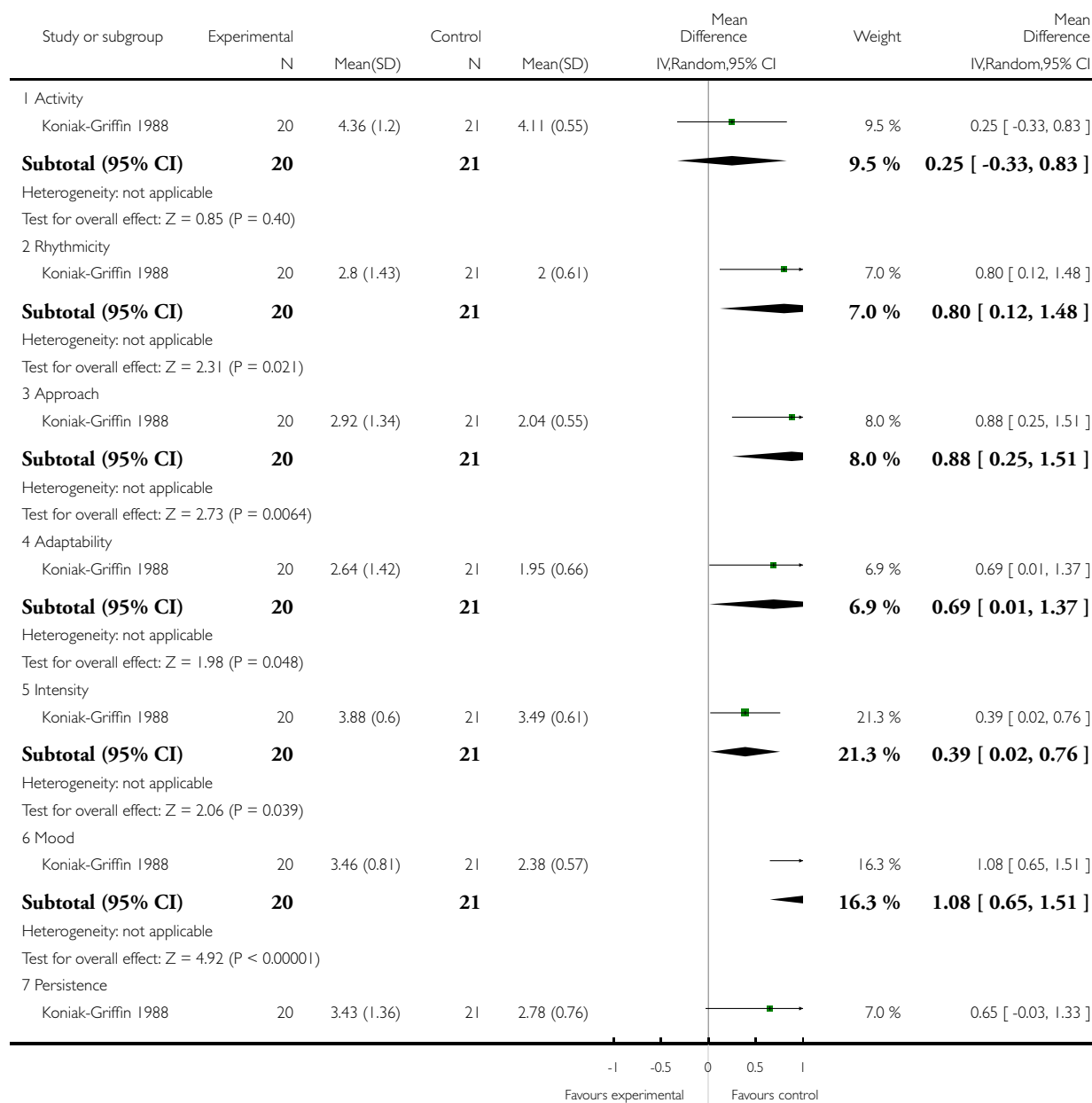


## Analysis 2.5. Comparison 2 Infant massage versus control - mental health and development, Outcome 5 Infant temperament questionnaire (revised RITQ (Carey)) follow-up 8 months.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

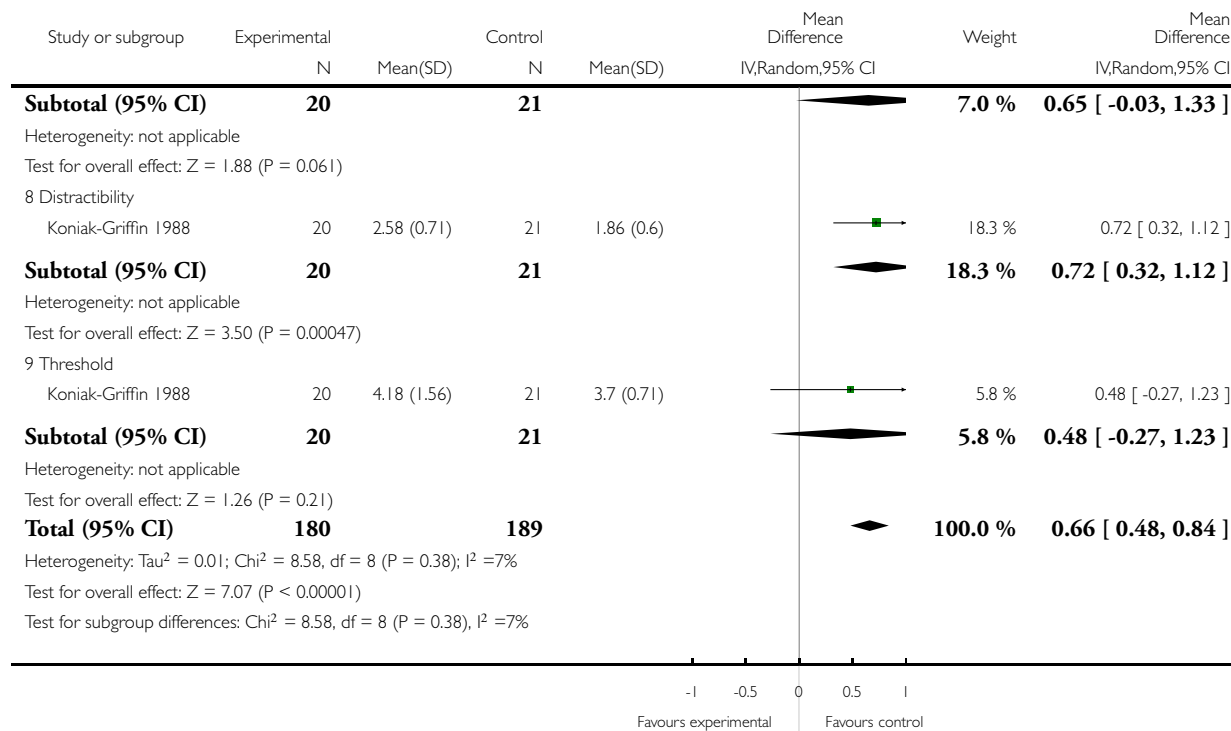
Comparison: 2 Infant massage versus control - mental health and development

Outcome: 5 Infant temperament questionnaire (revised RITQ (Carey)) follow-up 8 months



(Continued ...)

(... Continued)

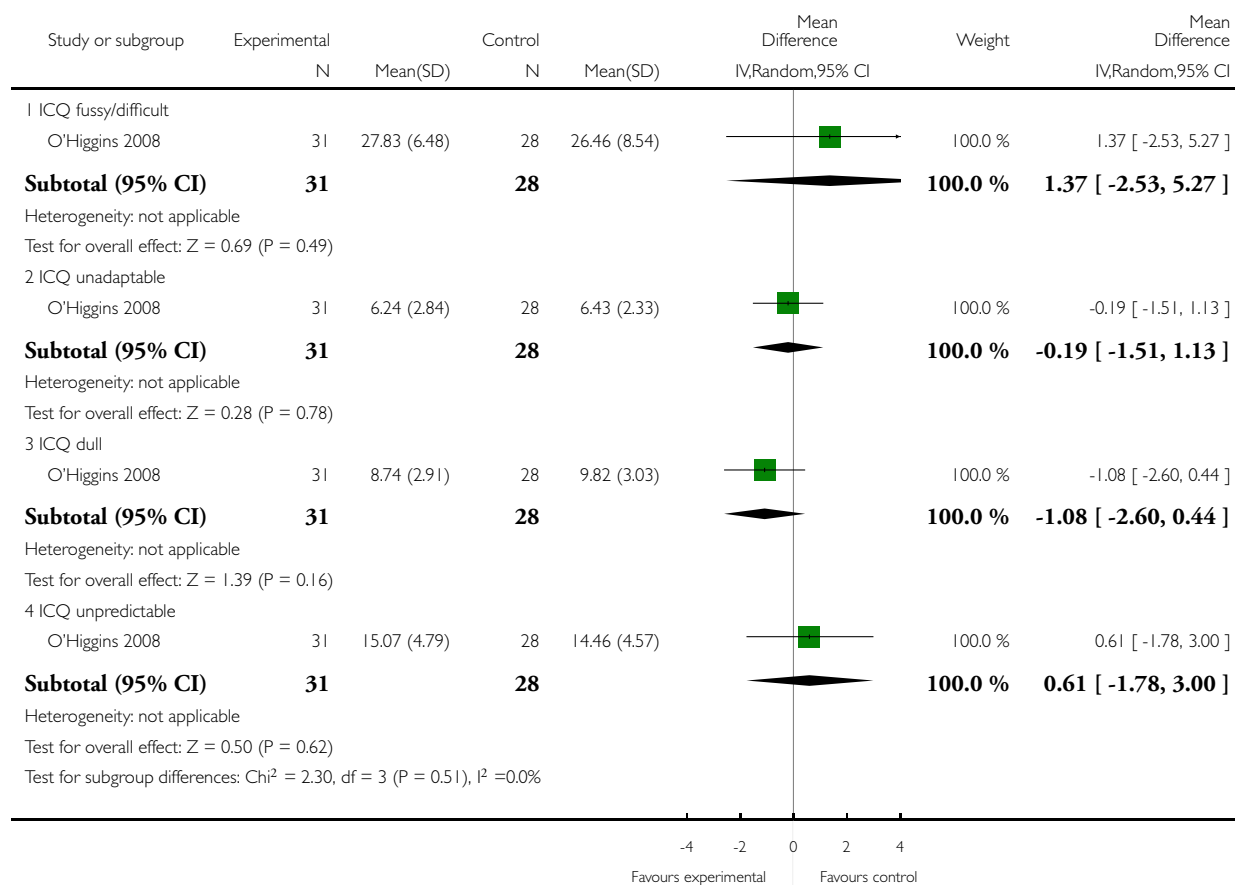


## Analysis 2.6. Comparison 2 Infant massage versus control - mental health and development, Outcome 6 Infant Care Questionnaire post-intervention.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 6 Infant Care Questionnaire post-intervention

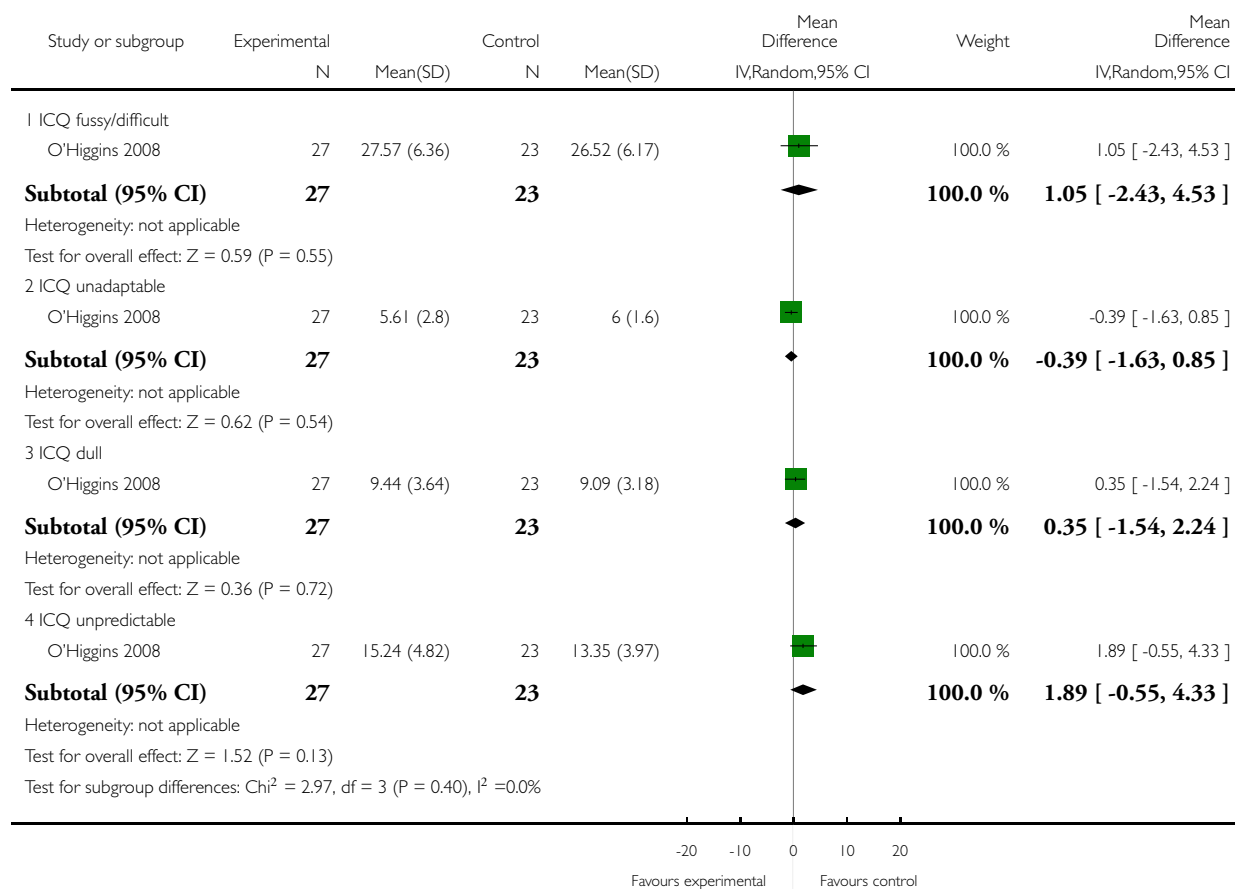


## Analysis 2.7. Comparison 2 Infant massage versus control - mental health and development, Outcome 7 Infant Care Questionnaire follow-up 1 year.

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 7 Infant Care Questionnaire follow-up 1 year

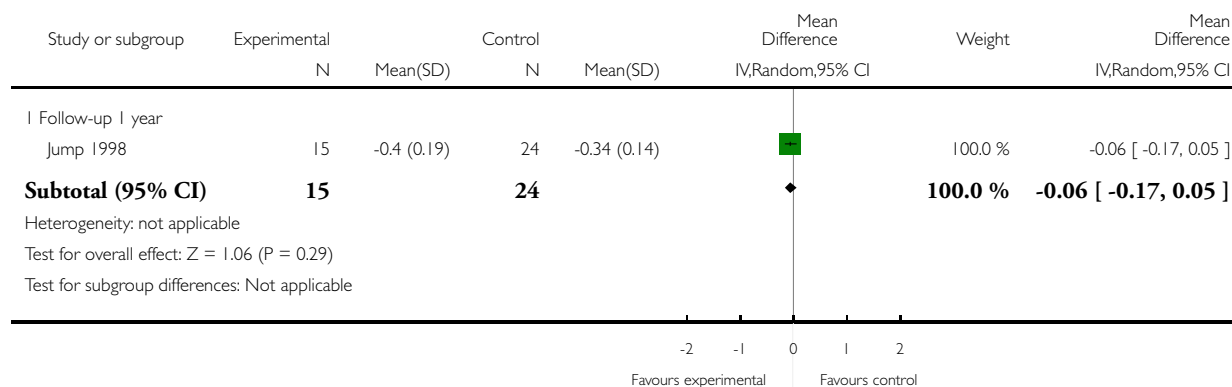


### Analysis 2.8. Comparison 2 Infant massage versus control - mental health and development, Outcome 8 Infant attachment (Q set).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 8 Infant attachment (Q set)

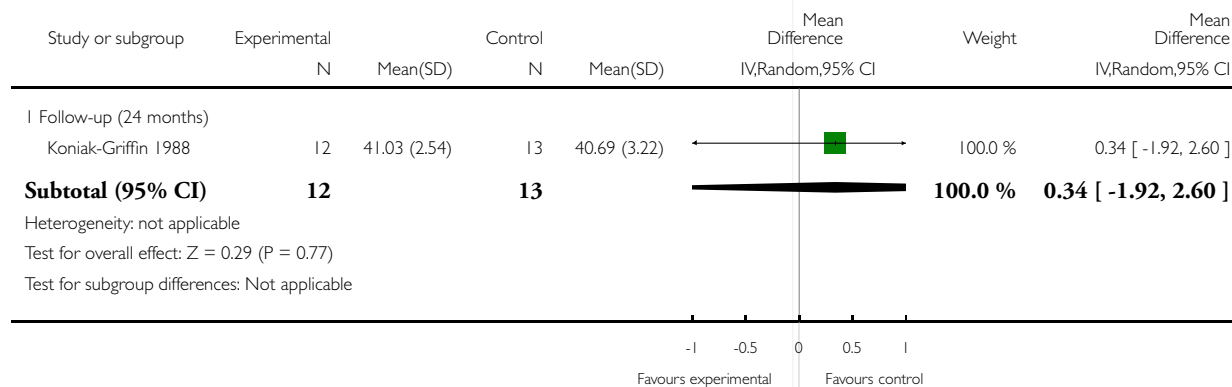


### Analysis 2.9. Comparison 2 Infant massage versus control - mental health and development, Outcome 9 Child behaviour (HOME).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 9 Child behaviour (HOME)



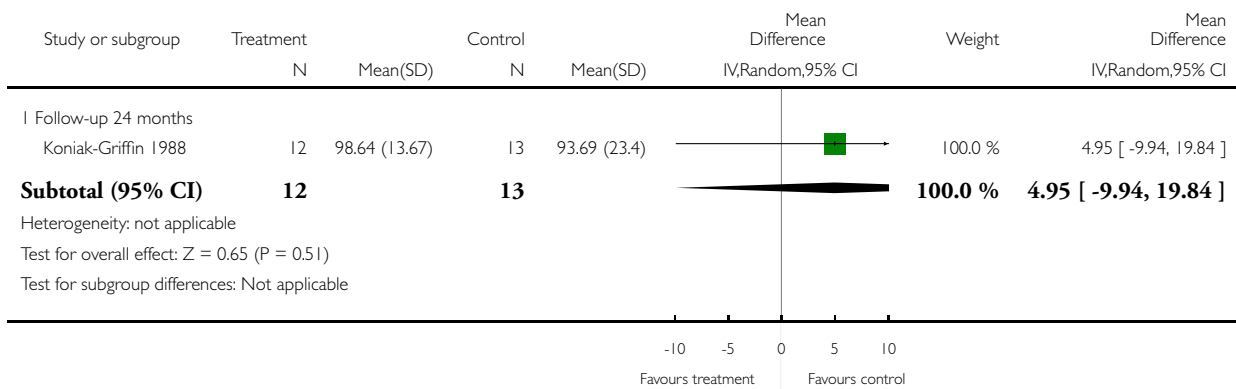


**Analysis 2.10. Comparison 2 Infant massage versus control - mental health and development, Outcome 10 Eyberg Child Behaviour Inventory (ECBI) - Intensity domain.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 10 Eyberg Child Behaviour Inventory (ECBI) - Intensity domain

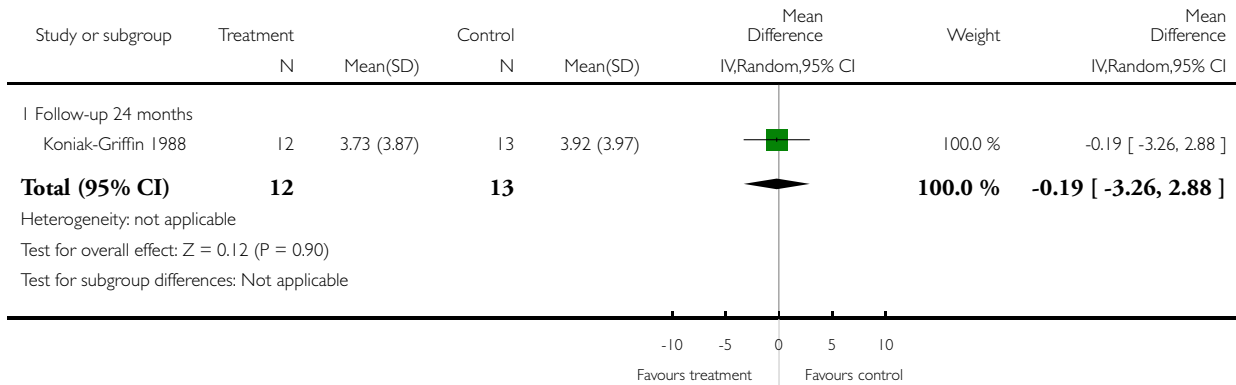


**Analysis 2.11. Comparison 2 Infant massage versus control - mental health and development, Outcome 11 Eyberg Child Behaviour Inventory (ECBI) - Problem domain.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 11 Eyberg Child Behaviour Inventory (ECBI) - Problem domain

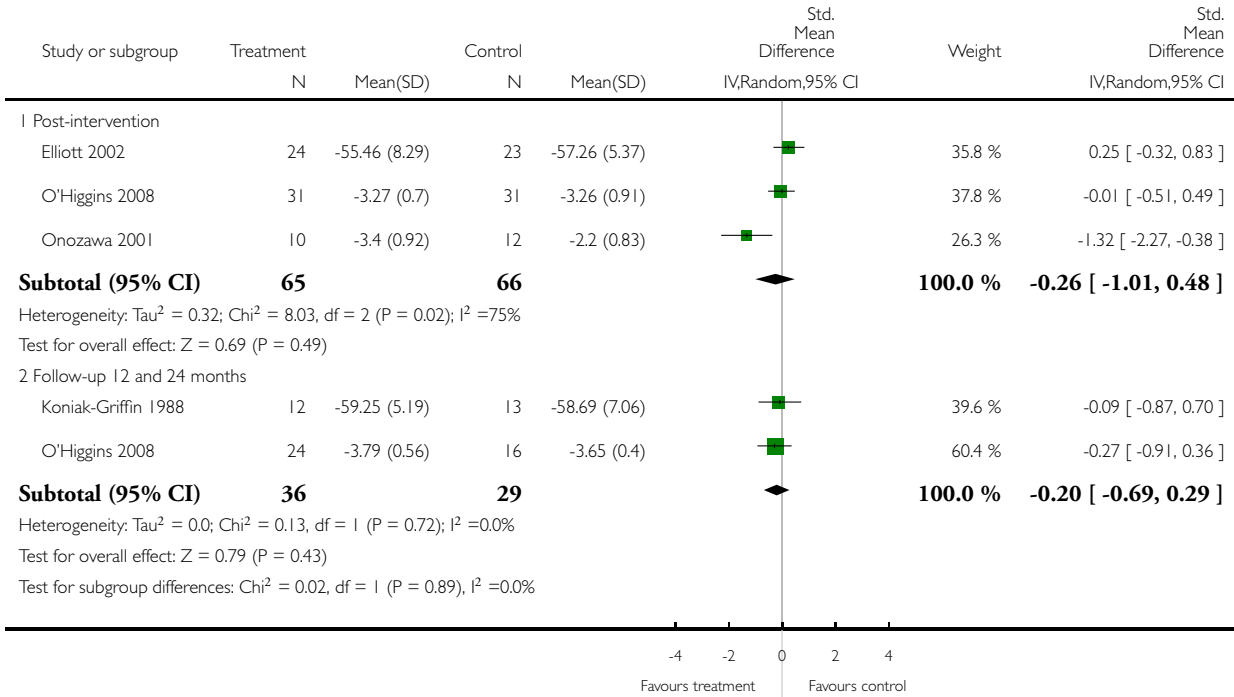


**Analysis 2.12. Comparison 2 Infant massage versus control - mental health and development, Outcome 12  
Mother and child interaction meta-analysis - Total NCATS and Murray Global.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 12 Mother and child interaction meta-analysis - Total NCATS and Murray Global

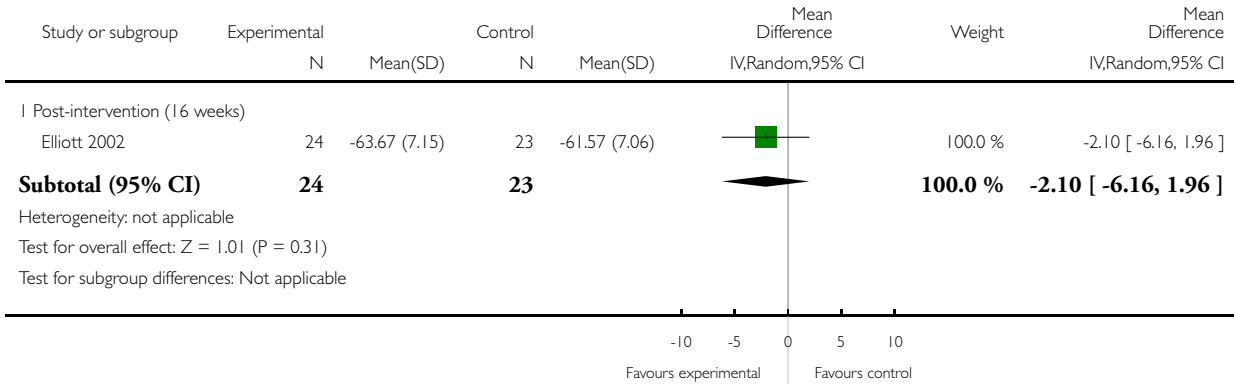


**Analysis 2.13. Comparison 2 Infant massage versus control - mental health and development, Outcome 13 Nursing Child Feeding Assessment Scale (NCAFS) - Total.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 13 Nursing Child Feeding Assessment Scale (NCAFS) - Total

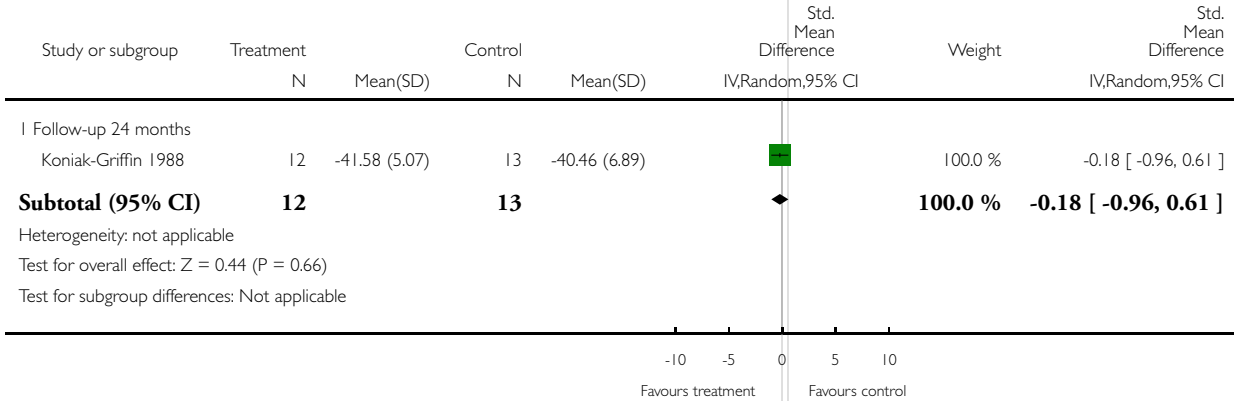


**Analysis 2.14. Comparison 2 Infant massage versus control - mental health and development, Outcome 14 Nursing Child Assessment Teaching Scale (NCATS) - Mother.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 14 Nursing Child Assessment Teaching Scale (NCATS) - Mother

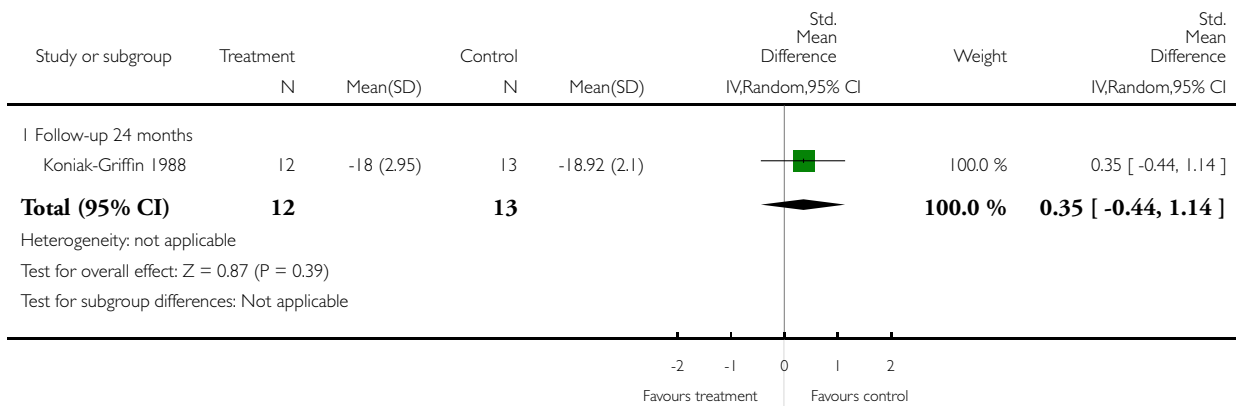


**Analysis 2.15. Comparison 2 Infant massage versus control - mental health and development, Outcome 15 Nursing Child Assessment Teaching Scale (NCATS) - Child.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 15 Nursing Child Assessment Teaching Scale (NCATS) - Child

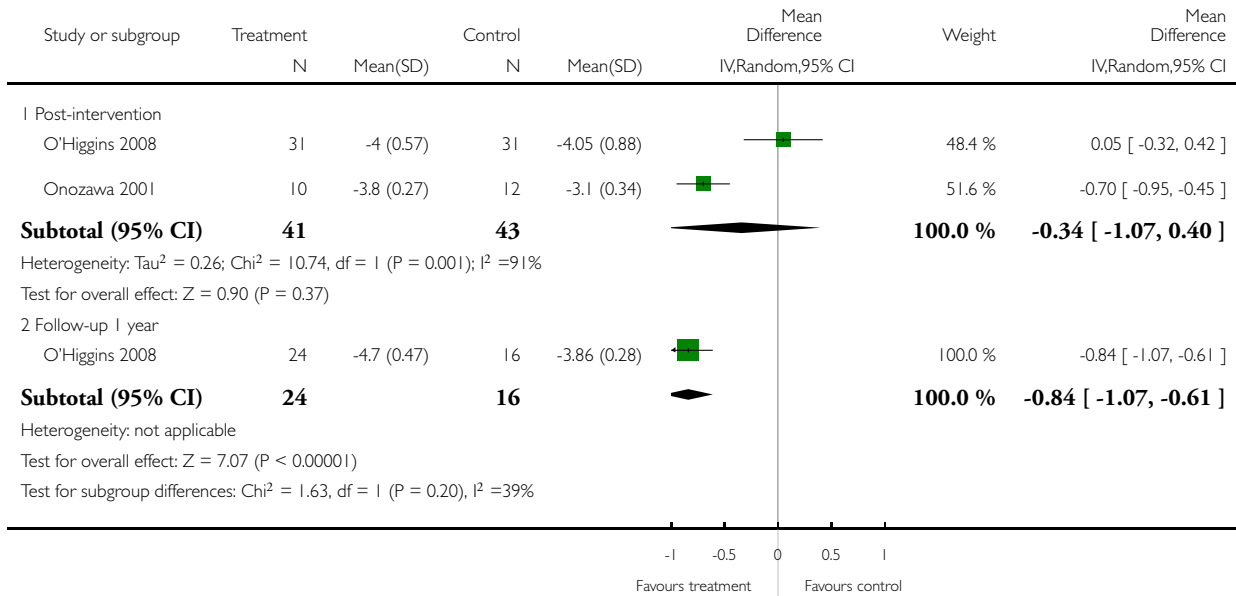


## Analysis 2.16. Comparison 2 Infant massage versus control - mental health and development, Outcome 16 Maternal sensitivity - warm to cold (Murray).

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 16 Maternal sensitivity - warm to cold (Murray)

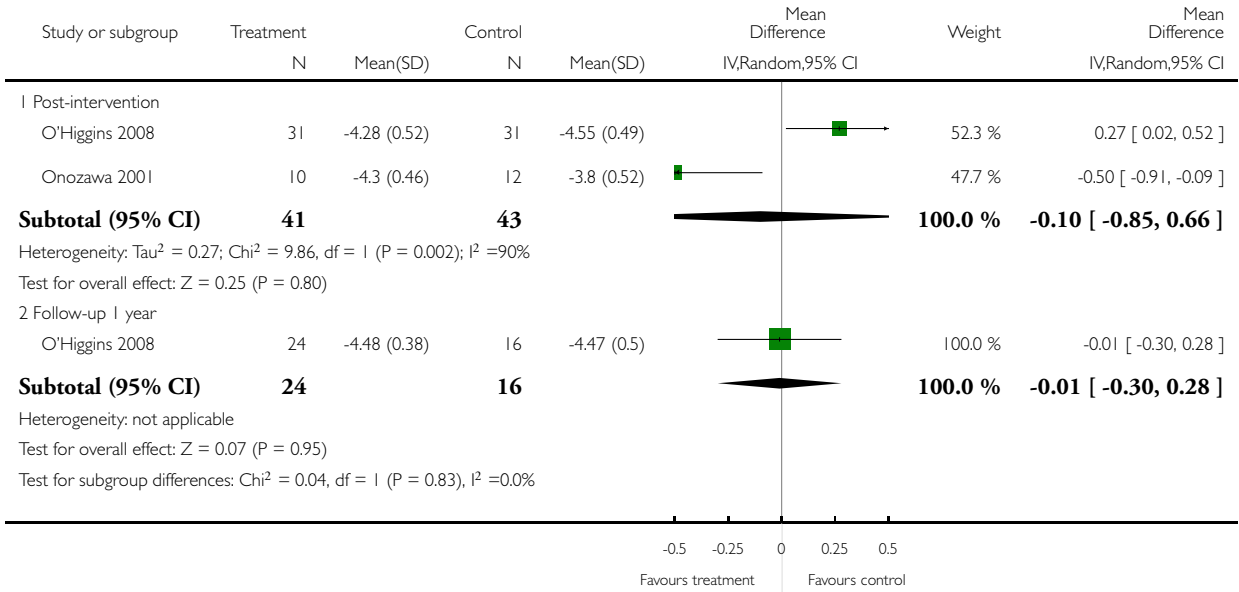


**Analysis 2.17. Comparison 2 Infant massage versus control - mental health and development, Outcome 17 Maternal sensitivity - non-intrusive to intrusive (Murray).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 17 Maternal sensitivity - non-intrusive to intrusive (Murray)

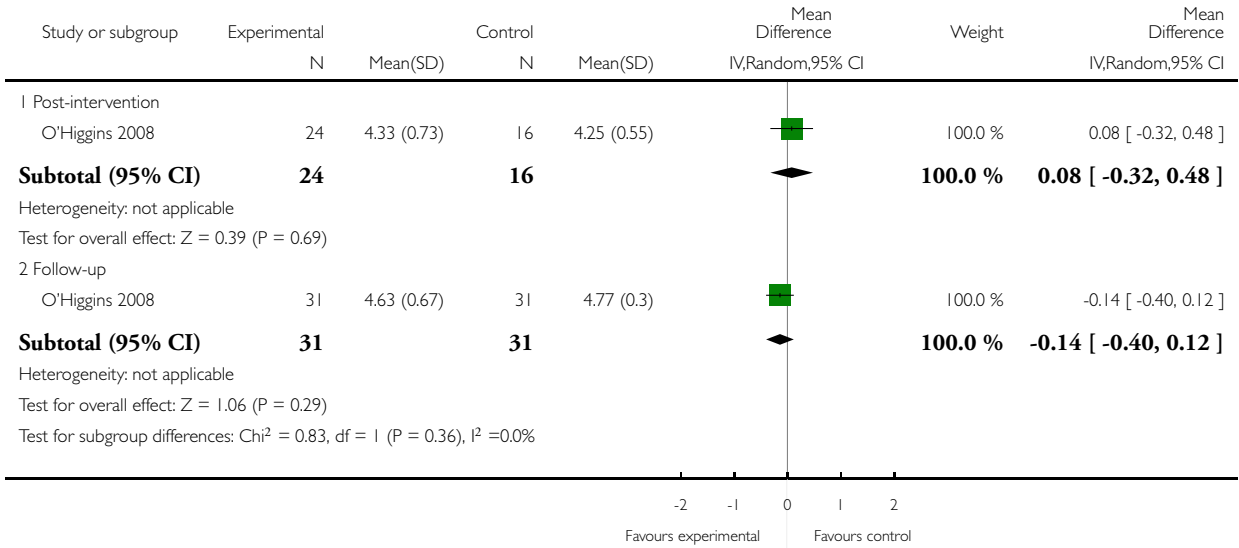


**Analysis 2.18. Comparison 2 Infant massage versus control - mental health and development, Outcome 18 Maternal sensitivity - remoteness (Murray).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 18 Maternal sensitivity - remoteness (Murray)



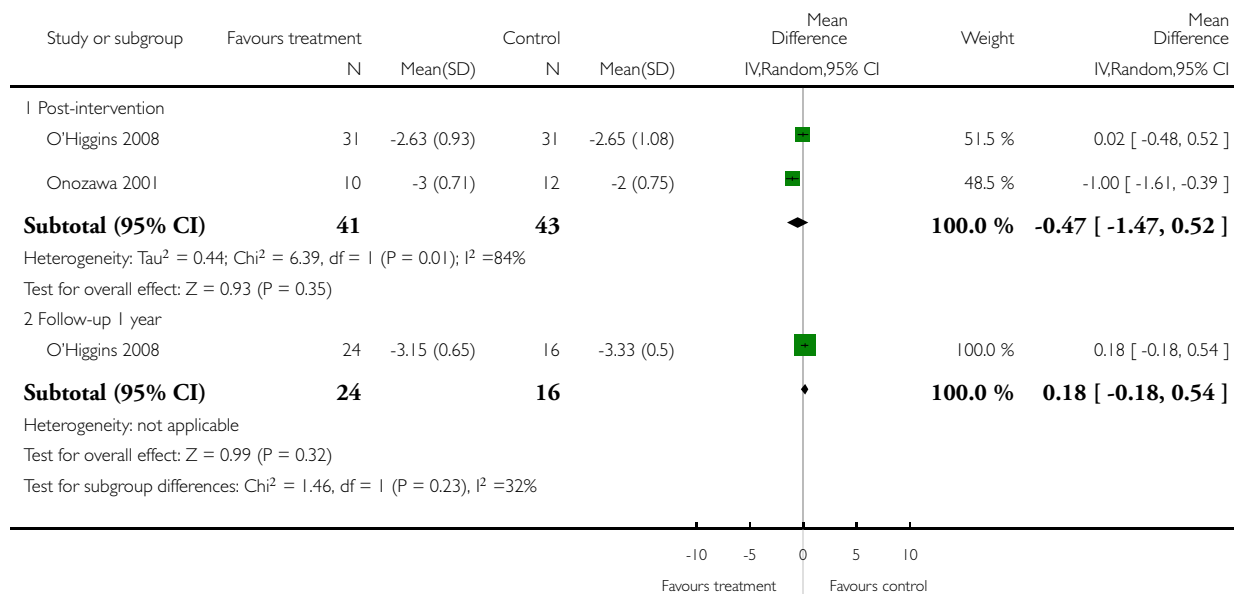


**Analysis 2.19. Comparison 2 Infant massage versus control - mental health and development, Outcome 19 Infant interactions - infant performance - attentive to non attentive (Murray).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 19 Infant interactions - infant performance - attentive to non attentive (Murray)

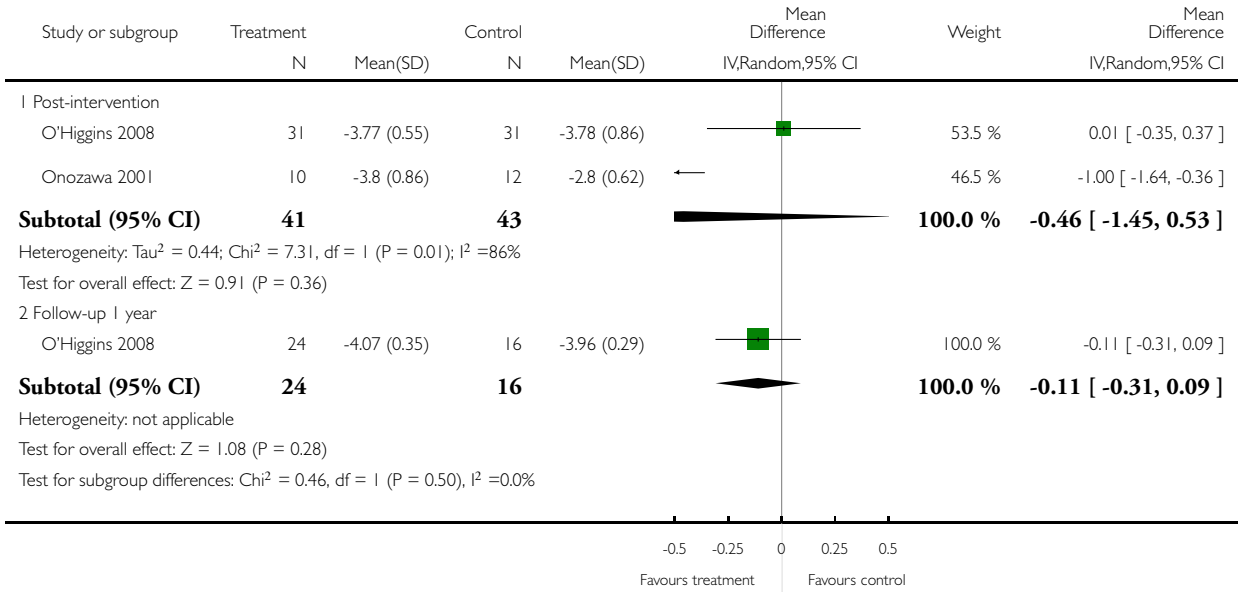


**Analysis 2.20. Comparison 2 Infant massage versus control - mental health and development, Outcome 20 Infant interactions - lively to inert (Murray).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 20 Infant interactions - lively to inert (Murray)

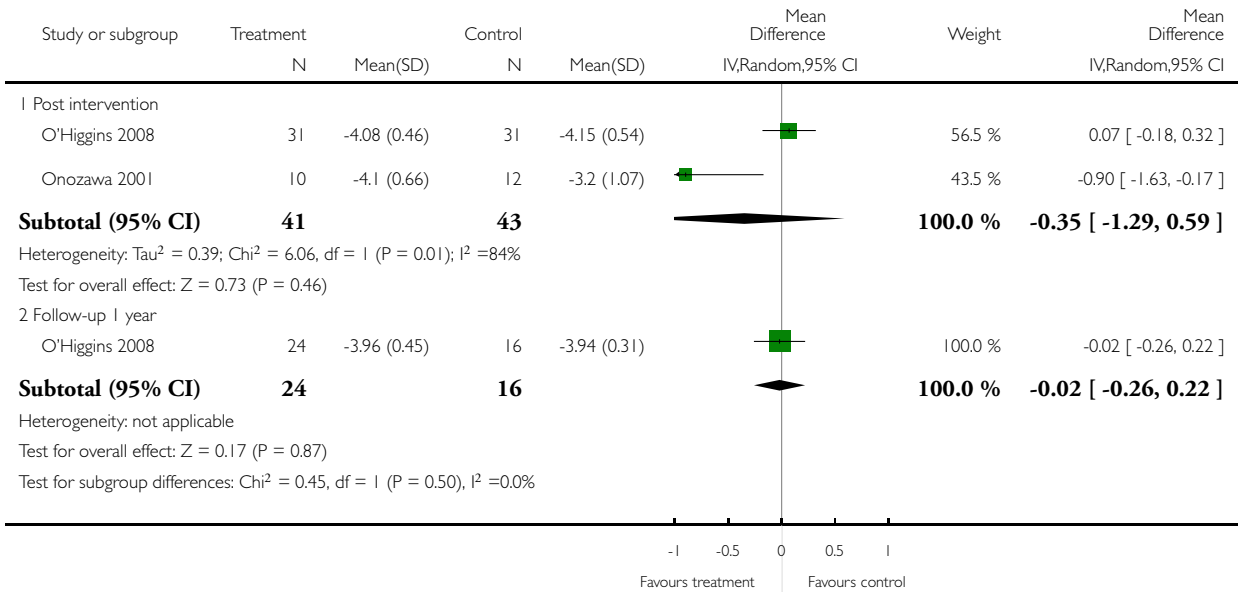


**Analysis 2.21. Comparison 2 Infant massage versus control - mental health and development, Outcome 21 Infant interactions - happy to distressed (Murray).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 21 Infant interactions - happy to distressed (Murray)

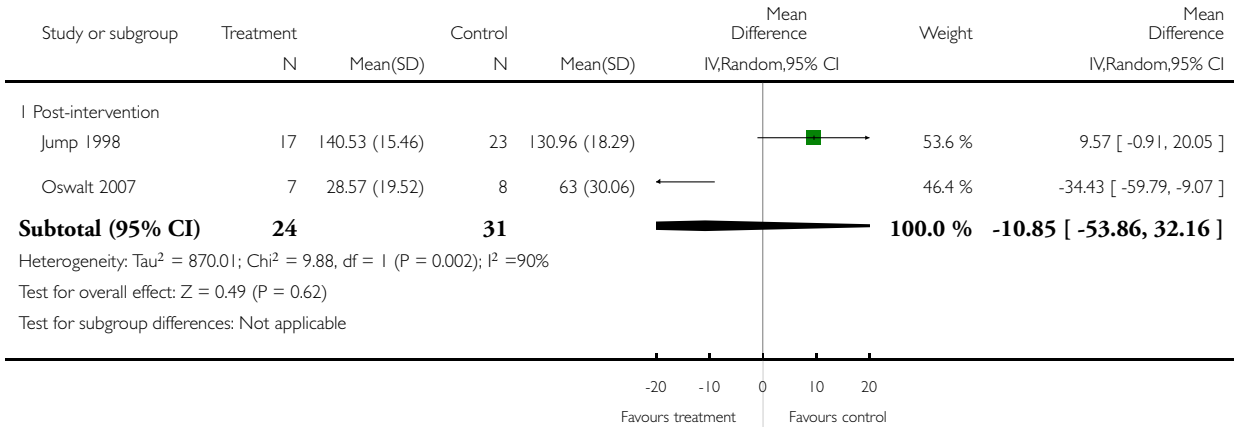


**Analysis 2.22. Comparison 2 Infant massage versus control - mental health and development, Outcome 22 Parenting stress (PSI Abidin) child characteristics subscale.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 22 Parenting stress (PSI Abidin) child characteristics subscale

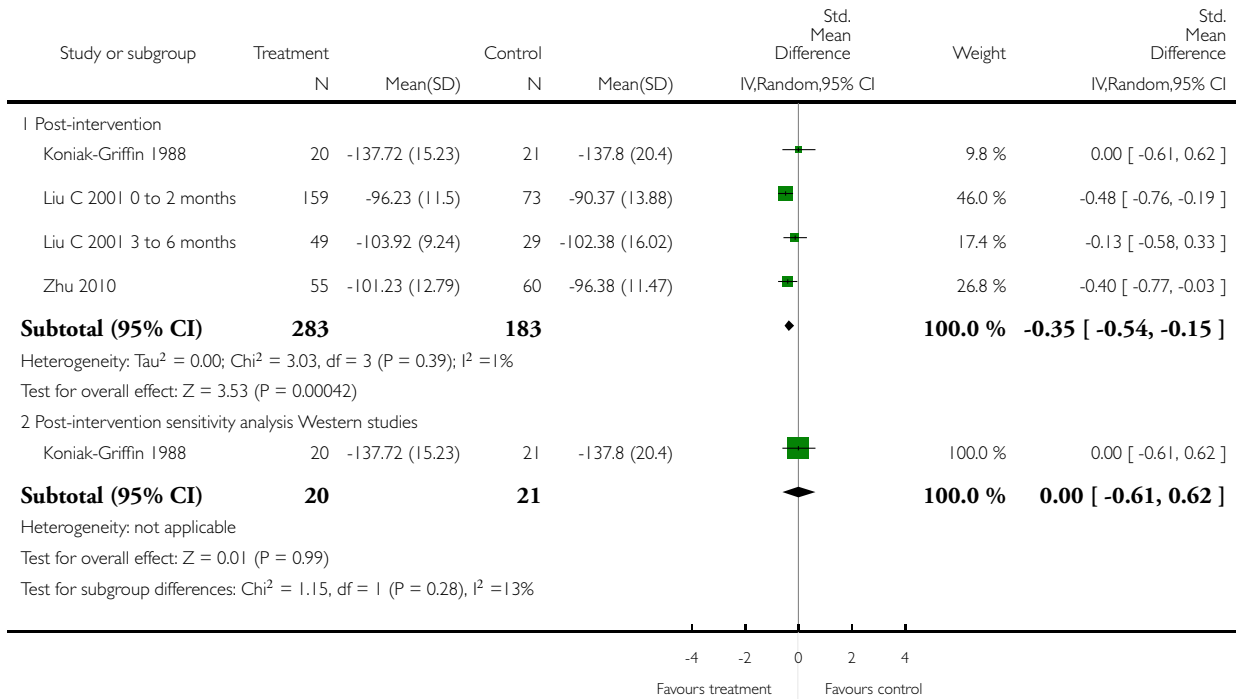


**Analysis 2.23. Comparison 2 Infant massage versus control - mental health and development, Outcome 23 Psychomotor Development Indices (PDI) meta-analysis post-intervention.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 23 Psychomotor Development Indices (PDI) meta-analysis post-intervention

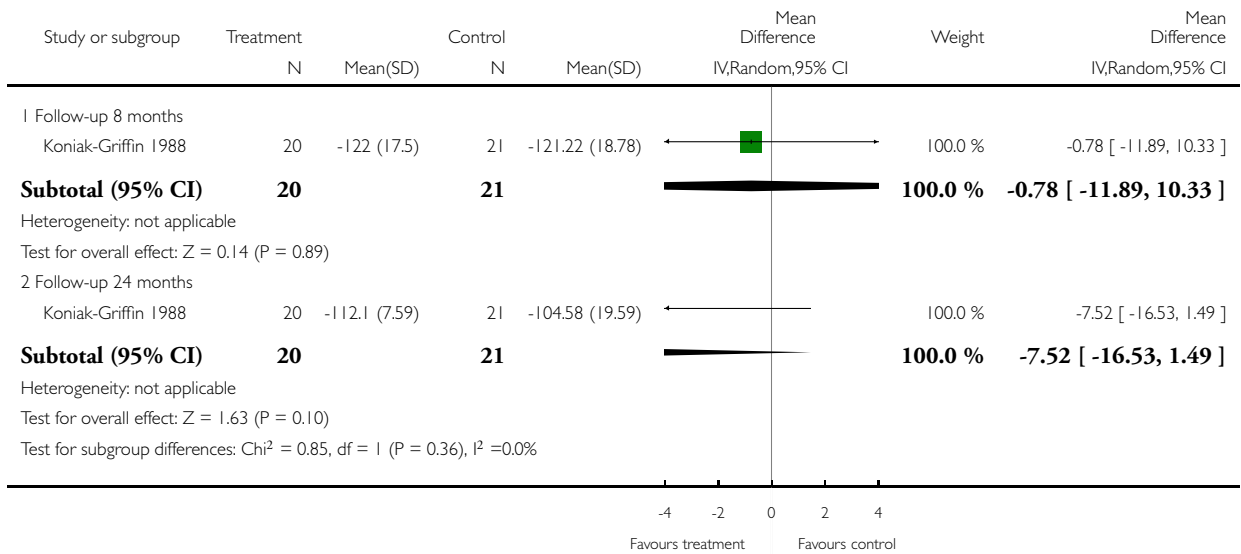


**Analysis 2.24. Comparison 2 Infant massage versus control - mental health and development, Outcome 24 Bayley Psychomotor Development Index (PDI) follow-up.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 24 Bayley Psychomotor Development Index (PDI) follow-up

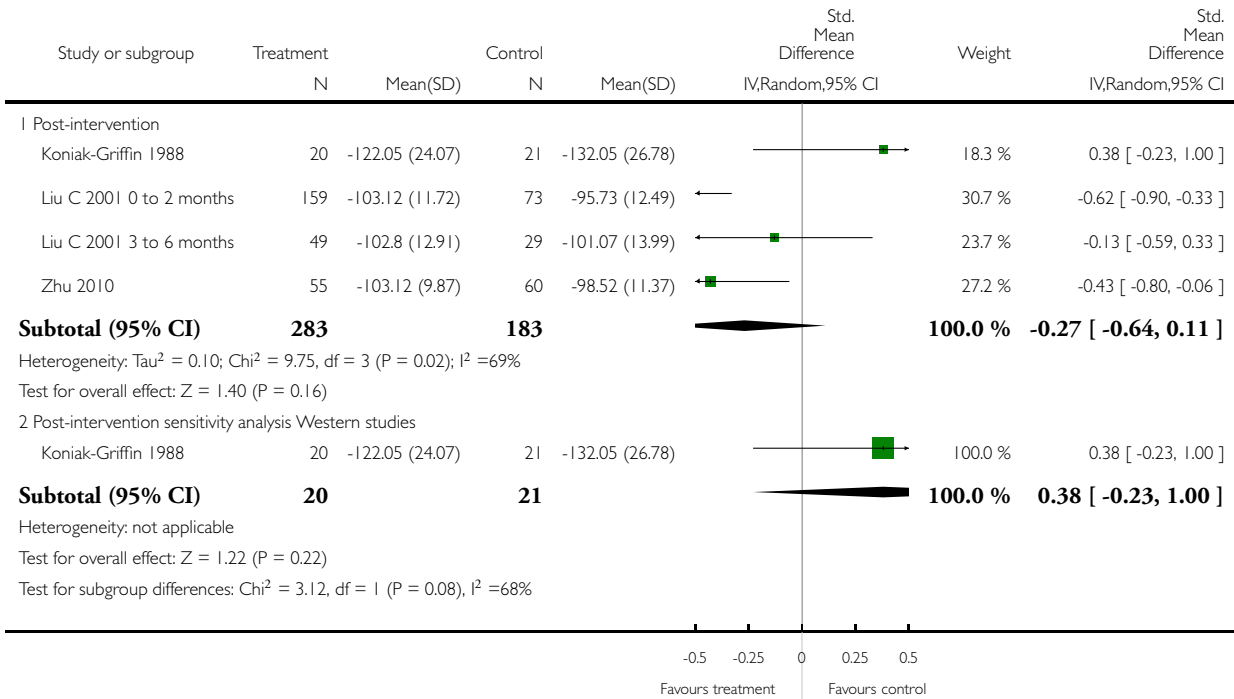


**Analysis 2.25. Comparison 2 Infant massage versus control - mental health and development, Outcome 25 Mental Development Indices (MDI) meta-analysis post-intervention.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 25 Mental Development Indices (MDI) meta-analysis post-intervention

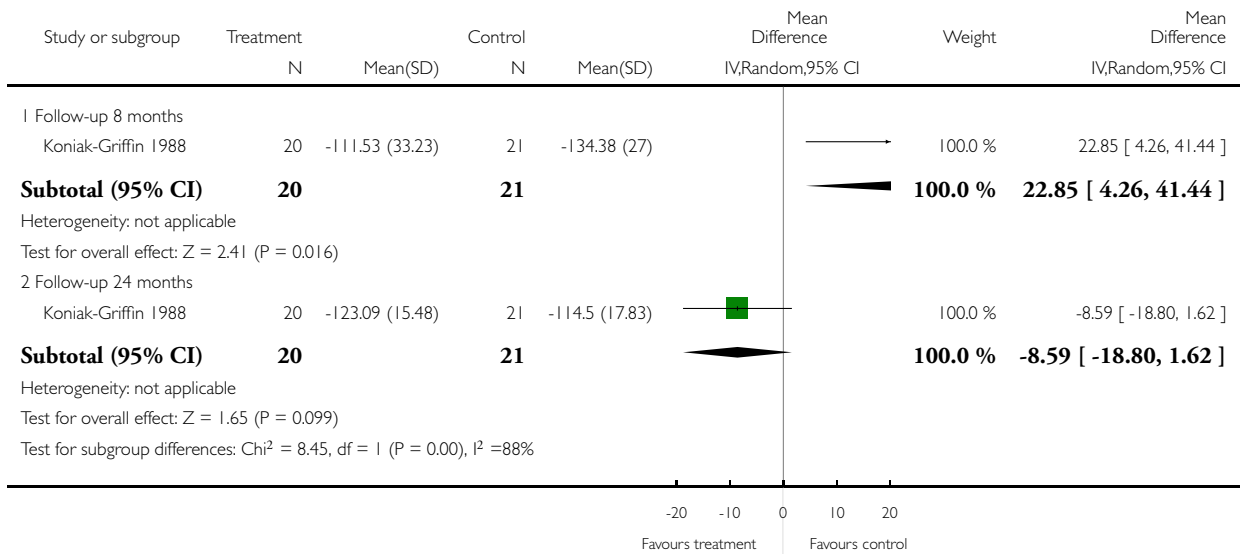


**Analysis 2.26. Comparison 2 Infant massage versus control - mental health and development, Outcome 26 Bayley Mental Development Index (MDI) follow-up.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 26 Bayley Mental Development Index (MDI) follow-up



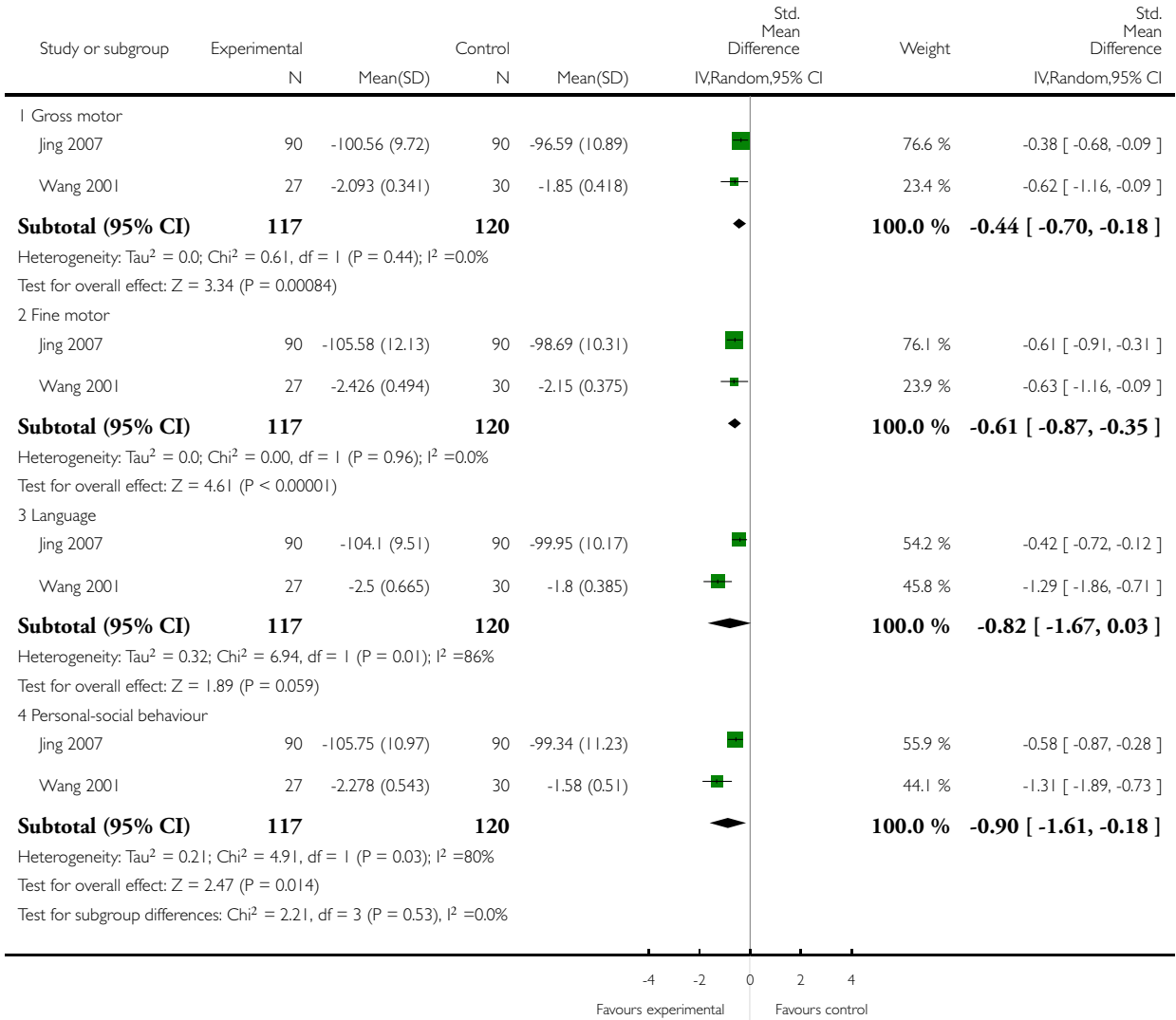


**Analysis 2.27. Comparison 2 Infant massage versus control - mental health and development, Outcome 27 Gessel/Capital meta-analysis (post intervention).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 27 Gessel/Capital meta-analysis (post intervention)



**Analysis 2.28. Comparison 2 Infant massage versus control - mental health and development, Outcome 28 Gessel Developmental Quotient (post intervention).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 28 Gessel Developmental Quotient (post intervention)

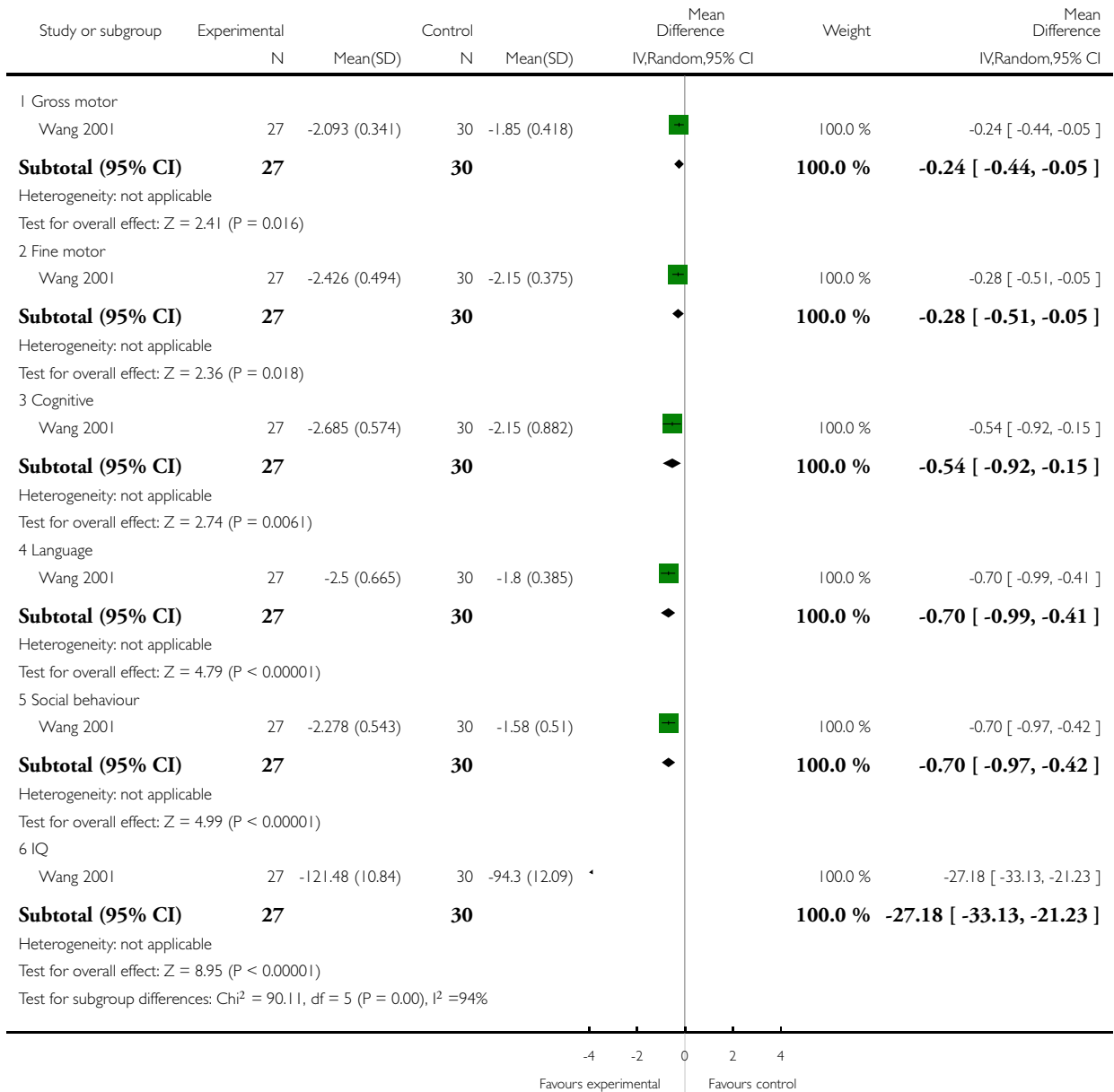
Study or subgroup	Experimental		Control		Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
	N	Mean(SD)	N	Mean(SD)			
1 Adaptive behaviour							
Jing 2007	90	-106.56 (9.11)	90	-99.49 (9.24) *		100.0 %	-7.07 [ -9.75, -4.39 ]
<b>Subtotal (95% CI)</b>	<b>90</b>		<b>90</b>			<b>100.0 %</b>	<b>-7.07 [ -9.75, -4.39 ]</b>
Heterogeneity: not applicable Test for overall effect: Z = 5.17 (P < 0.00001)							
2 Gross motor							
Jing 2007	90	-100.56 (9.72)	90	-96.59 (10.89) *		100.0 %	-3.97 [ -6.99, -0.95 ]
<b>Subtotal (95% CI)</b>	<b>90</b>		<b>90</b>			<b>100.0 %</b>	<b>-3.97 [ -6.99, -0.95 ]</b>
Heterogeneity: not applicable Test for overall effect: Z = 2.58 (P = 0.0099)							
3 Fine motor							
Jing 2007	90	-105.58 (12.13)	90	-98.69 (10.31) *		100.0 %	-6.89 [ -10.18, -3.60 ]
<b>Subtotal (95% CI)</b>	<b>90</b>		<b>90</b>			<b>100.0 %</b>	<b>-6.89 [ -10.18, -3.60 ]</b>
Heterogeneity: not applicable Test for overall effect: Z = 4.11 (P = 0.000040)							
4 Language							
Jing 2007	90	-104.1 (9.51)	90	-99.95 (10.17) *		100.0 %	-4.15 [ -7.03, -1.27 ]
<b>Subtotal (95% CI)</b>	<b>90</b>		<b>90</b>			<b>100.0 %</b>	<b>-4.15 [ -7.03, -1.27 ]</b>
Heterogeneity: not applicable Test for overall effect: Z = 2.83 (P = 0.0047)							
5 Personal-social behaviour							
Jing 2007	90	-105.75 (10.97)	90	-99.34 (11.23) *		100.0 %	-6.41 [ -9.65, -3.17 ]
<b>Subtotal (95% CI)</b>	<b>90</b>		<b>90</b>			<b>100.0 %</b>	<b>-6.41 [ -9.65, -3.17 ]</b>
Heterogeneity: not applicable Test for overall effect: Z = 3.87 (P = 0.00011) Test for subgroup differences: Chi <sup>2</sup> = 4.07, df = 4 (P = 0.40), I <sup>2</sup> = 2%							

**Analysis 2.29. Comparison 2 Infant massage versus control - mental health and development, Outcome 29 Capital institute Mental Checklist (post intervention).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 29 Capital institute Mental Checklist (post intervention)

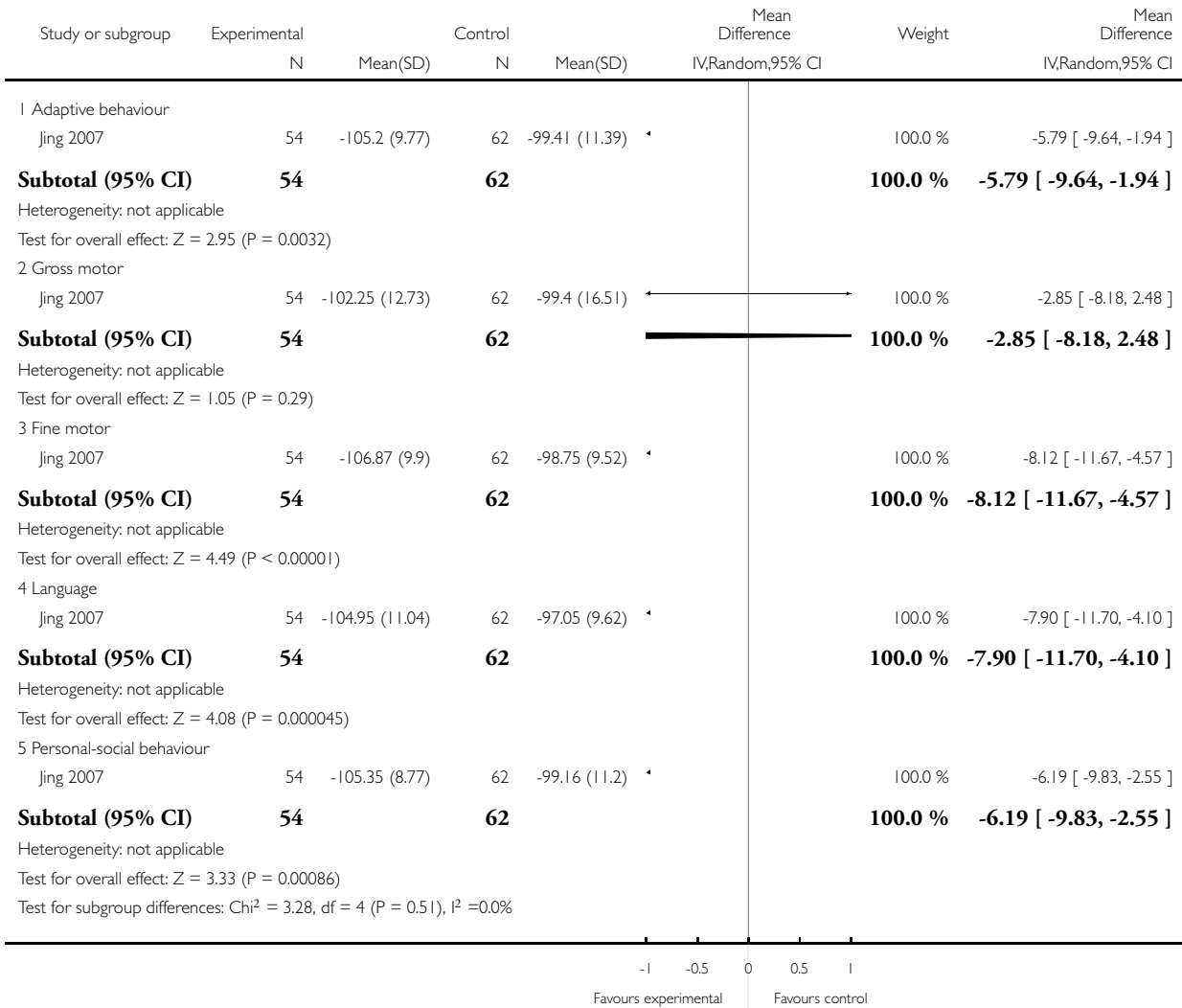


**Analysis 2.30. Comparison 2 Infant massage versus control - mental health and development, Outcome 30 Gessel Developmental Quotient (follow-up 6 months).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 30 Gessel Developmental Quotient (follow-up 6 months)

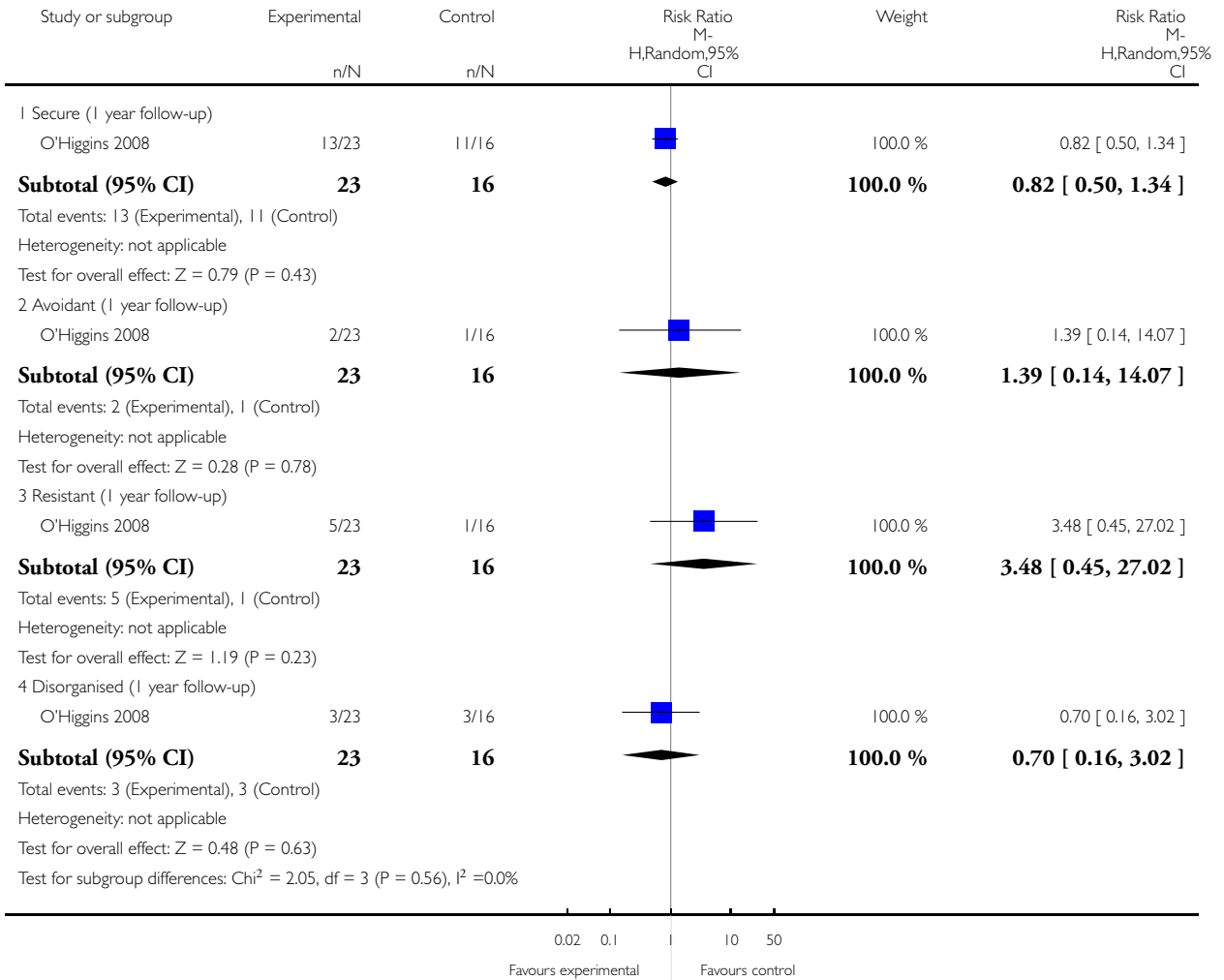


**Analysis 2.31. Comparison 2 Infant massage versus control - mental health and development, Outcome 31 Attachment patterns (strange situation procedure).**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 31 Attachment patterns (strange situation procedure)

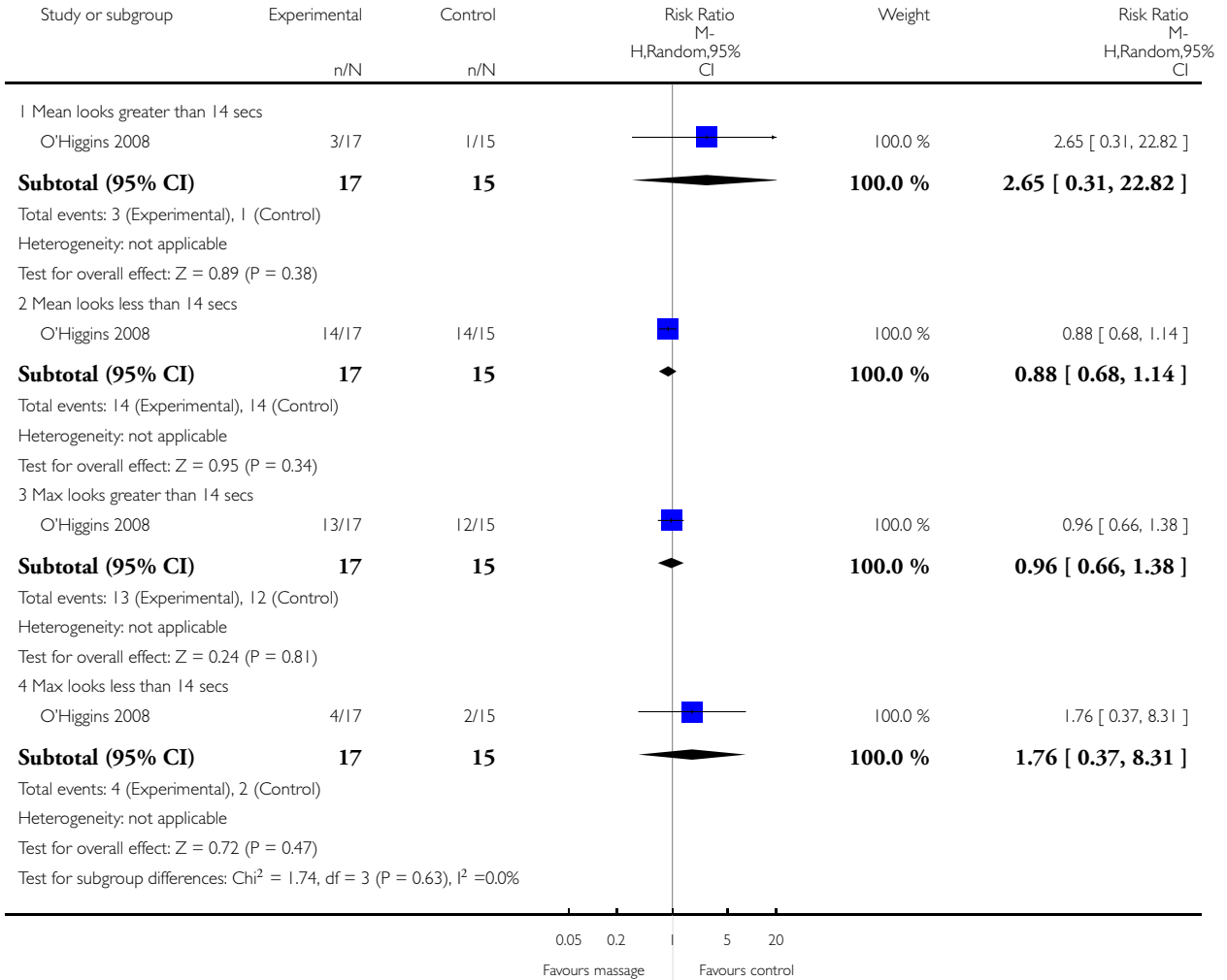


**Analysis 2.32. Comparison 2 Infant massage versus control - mental health and development, Outcome 32 Distractibility (toy) follow-up 1 year.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 32 Distractibility (toy) follow-up 1 year

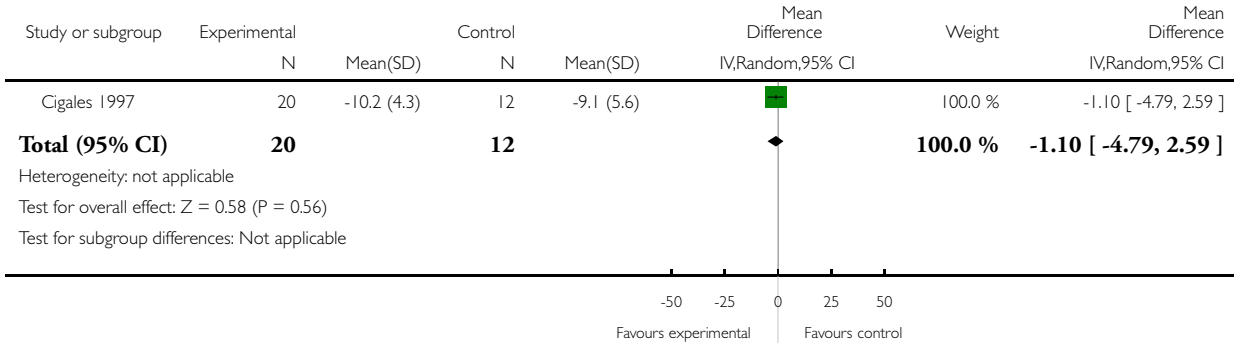


**Analysis 2.33. Comparison 2 Infant massage versus control - mental health and development, Outcome 33 Habituation.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 33 Habituation

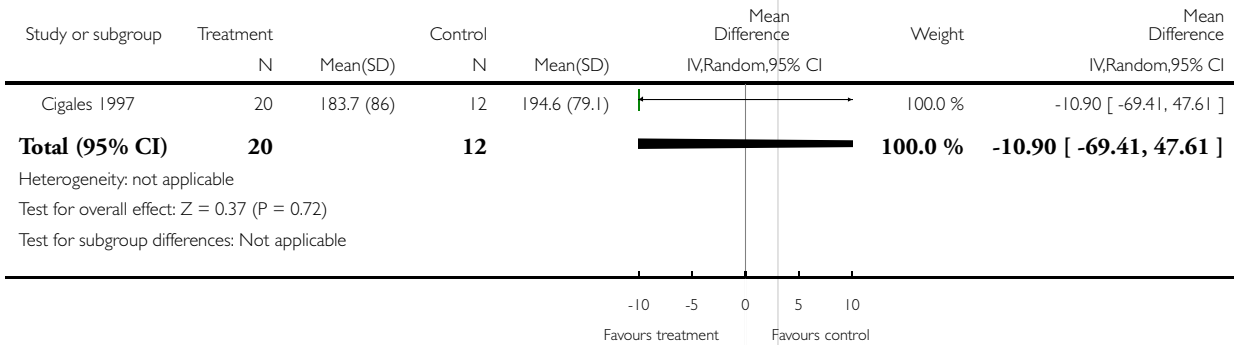


**Analysis 2.34. Comparison 2 Infant massage versus control - mental health and development, Outcome 34 Seconds to habituation.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 34 Seconds to habituation

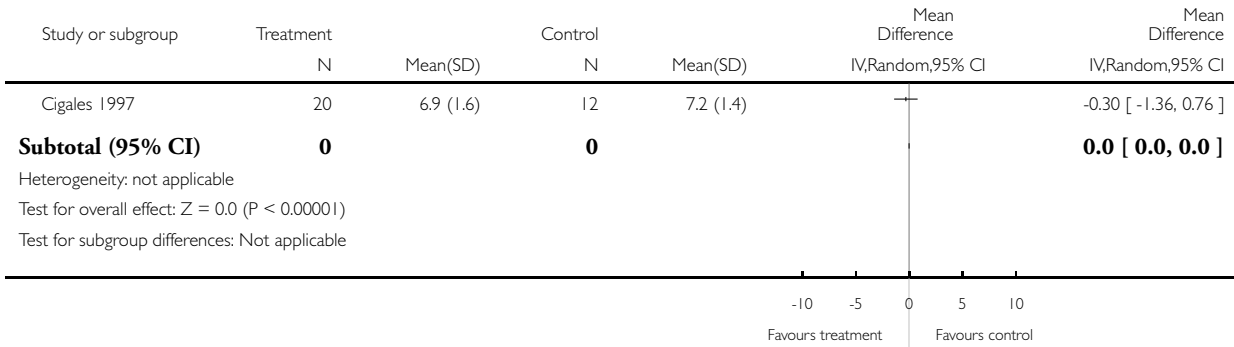


**Analysis 2.35. Comparison 2 Infant massage versus control - mental health and development, Outcome 35  
Trials to habituation.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 35 Trials to habituation

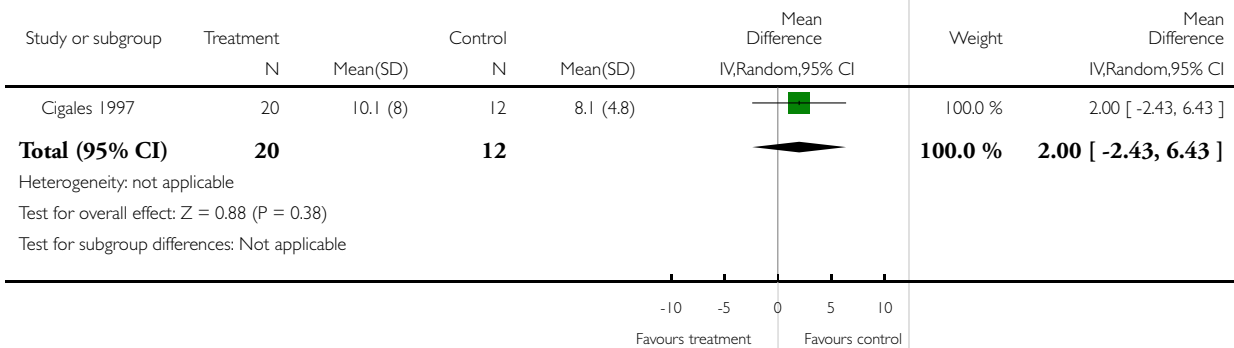


**Analysis 2.36. Comparison 2 Infant massage versus control - mental health and development, Outcome 36  
Post habituation.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 36 Post habituation



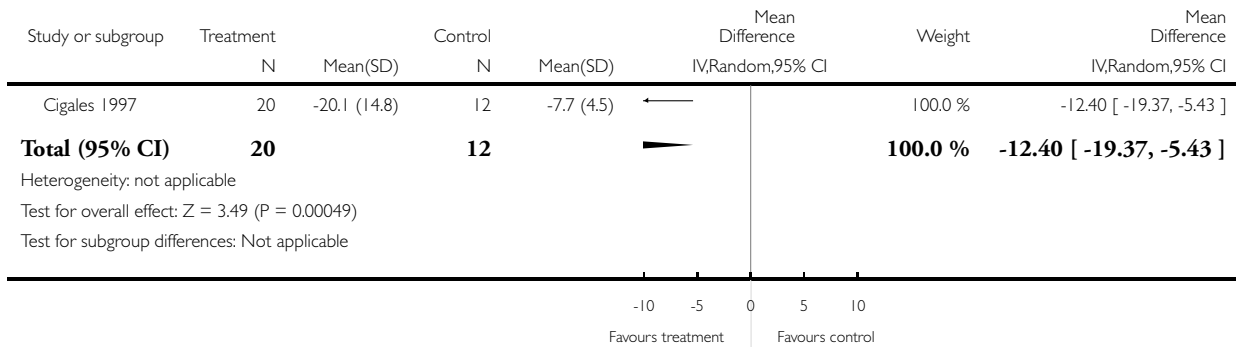


**Analysis 2.37. Comparison 2 Infant massage versus control - mental health and development, Outcome 37 Habituation test.**

Review: Massage for promoting mental and physical health in typically developing infants under the age of six months

Comparison: 2 Infant massage versus control - mental health and development

Outcome: 37 Habituation test



**ADDITIONAL TABLES**

Table 1. Study investigators' analyses: comparison of physical development

	Survey time	Height	Weight	Head	Chest	Comment
<a href="#">Liu C 2001 0 to 2 months</a> ; <a href="#">Liu C 2001 3 to 6 months</a>	4 months of age (1 month Post-intervention)	t = 0.854; P = 0.396	t = 1.120; P = 0.226	t = -0.343; P = 0.732	t = 0.995; P = 0.322	Through a six-month vertical survey of the growth of all n = 310 (that is, all participants from both <a href="#">Liu C 2001 0 to 2 months</a> ; <a href="#">Liu C 2001 3 to 6 months</a> ) the infant participants over 0-6 months, it was shown that the weight and the chest circumference of the infants who received the mas-

**Table 1. Study investigators' analyses: comparison of physical development** (Continued)

							sage developed better than the control group. There was a significant difference between infants of the two groups by the six months. Height and head circumference were not significantly different * Significantly different
	6 months of age (3 months Post-intervention)	t = 1.763; P = 0.081	t = 2.295; *P = 0.025	t = 0.411; P = 0.682	t = 2.659; *P = 0.010		
Maimaiti 2007	n/a	n/a	n/a	n/a	n/a		Outcome assessments at Post-intervention on weight, length and head circumference were presented using a $\chi^2$ sided test and were significantly different between massage and control group (P > 0.05)

**Table 2. Sleep habits**

Study ID	Intervention	Good	Medium	Not good	Control	Good	Medium	Not good	Statistical significance X2 P
Liu C 2001 0 to 2 months	n = 159	136	23	0	n = 73	49	20	4	X2 = 15.353 P = 0.0000 (statistically significant between massage and control)
Liu C 2001 3 to 6 months	n = 41	41	7	1	n = 29	21	7	1	X2 = 1.417 P = > 0.10 (not statistically significant between massage and control)

**Table 3. Other developmental measures**

Study ID	Outcome measure (Post-intervention)	Intervention	Control	Statistical tests X <sup>2</sup> P
Maimaiti 2007	Rise from prone 0 degrees	6	71	X <sup>2</sup> = 4.212; P = < 0.05 Statistically significant between intervention and control.
	Rise from prone 45 degrees	61	23	
	Rise from prone 90 degrees	33	6	
	Sight tracking 30cm	19	41	X <sup>2</sup> = 30.11; P = < 0.05 Statistically significant between intervention and control.
	Sight tracking 50cm	42	39	
	Sight tracking 100cm	39	20	
	Auditory tracking Can do	91	86	X <sup>2</sup> = 4.735; P = < 0.05 Statistically significant between intervention and control.
	Auditory tracking Cannot do	9	14	
	Smiling for testers Can do	34	19	X <sup>2</sup> = 4.568; P = 0.05 Statistically significant between intervention and control.
	Smiling for testers Cannot do	66	81	

## APPENDICES

### Appendix I. Original search strategies

The following strategy was used to search CENTRAL:

- #1 MASSAGE (Mesh)
- #2 Massage next therap\*
- #3 Therapeutic next touch
- #4 THERAPEUTIC TOUCH (Mesh)
- #5 TOUCH
- #6 Tactile next stimulation
- #7 #1 or #2 or #3 or #4 or #5 or #6
- #8 infant\* or baby or babies
- #9 #7 and #8

The strategy used for MEDLINE, AMED and CINAHL was:

- #1 exp MASSAGE/ or massage.mp.
- #2 (massage adj therap\$).mp
- #3 (therapeutic adj touch).mp
- #4 exp TOUCH/ or exp THERAPEUTIC TOUCH/ or touch.mp.
- #5 (tactile adj stimulation).mp.

#6 1 or 2 or 3 or 4 or 5

#7 (infant\$ or baby or babies).mp.

#8 6 and 7

The strategy used for EMBASE was:

#1 exp MASSAGE/ or massage.mp.

#2 (massage adj therap\$).mp

#3 (therapeutic adj touch).mp

#4 exp TOUCH/ or touch.mp.

#5 (tactile adj stimulation).mp.

#6 1 or 2 or 3 or 4 or 5

#7 (infant\$ or baby or babies).mp.

#8 6 and 7 (1415)

The strategy used for LILACS was:

massage or massage therapy or massage therapies [Words] or therapeutic touch or touch or tactile stimulation [Words] and infant or infants or baby or babies [Words]

The strategy used for PsycINFO was:

#10 (infant\* or baby or babies) and (#1 or #2 or #3 or #4 or #5 or #6 or #7)

#9 infant\* or baby or babies

#8 #1 or #2 or #3 or #4 or #5 or #6 or #7

#7 tactile adj stimulation

#6 therapeutic touch

#5 TOUCH

#4 therapeutic adj touch\*

#3 massage adj therap\*

#2 massage

#1 MASSAGE

The strategy used for the National Research Register was:

#1 MASSAGE (Mesh)

#2 Massage next therap\*

#3 Therapeutic next touch

#4 THERAPEUTIC TOUCH (Mesh)

#5 TOUCH

#6 Tactile next stimulation

#7 #1 or #2 or #3 or #4 or #5 or #6

#8 infant\* or baby or babies

#9 #7 and #8

Dissertation Abstracts, ClinicalTrials.gov, Cochrane Neonatal Review Group specialised register and the Chinese databases were searched using the terms:

Infant or infants or baby or babies AND massage

## Appendix 2. Original search dates

Cochrane Central Register of Controlled Trials (CENTRAL) 2005 (Issue 3)

MEDLINE (1970 to August 2005)

PsycINFO (1970 to August 2005)

CINAHL (1982 to August 2005)

EMBASE (1980 to August 2005)

Dissertation Abstracts (1981 to August 2005)

AMED (Alternative and Complementary Medicine) (1985 to August 2005)

LILACS (Latin American & Caribbean Health Sciences Literature) (1982 to August 2005)

The National Research Register (2005) Issue 3

Clinicaltrials.gov (1966 to 2005)

Cochrane Neonatal Review Group specialised register (1966 to 2005)  
 VC undertook a search of the following Chinese database(s):  
 Chinese Scientific Journal Database (Jan 89 - Oct 05)  
 Traditional Chinese Medicine Database (Jan 84 - Sept 05)  
 Chinese BioMedical Database (Jan 89 - Oct 05)  
 China Academic Journal Full Text Database (Jan 94 - Oct 05)  
 China Proceedings of Conference Databases (Jan 99 - Oct 05)  
 China Doctorate/Master Dissertations Full Text Databases (Jan 99 - Oct 05)

### Appendix 3. Results of the updated searches

Database searched	Date of search	Issue searched	Number of hits	Date range of search
CENTRAL	14.05.2010	2011(2)	115	2005-2010
MEDLINE	12.05.2010	1950 to May Week 1 2011	478	2005-2010
EMBASE	17.05 2010	1980 to Week 19 2010	384	2005-2010
CINAHL	14.05.2010	1937 to current	368	2005-2010
PsycINFO	15.05.2010	1887 to current	299	2005-2010
Maternity and Infant Care	18.05.2010	1971 to May 2010	505	All years searched
LILACS	19.05.2010	all available years	30	2005-2010

Database searched	Date of search	Issue searched	Number of hits	Date range of search
CENTRAL	20.06.2011	2011(3)	21	Records added since May 2010
MEDLINE	20.06.2011	1948 to June Week 2 2011	121	Records added since May 2010
EMBASE	20.06.2011	1980 to 2011	159	Records added since May 2010
CINAHL	20.06.2011	1937 to current	111	Records added since May 2010
PsycINFO	20.06.2011	1887 to current	69	Records added since May 2010
Maternity and Infant Care	20.06.2011	1971 to May 2011	37	Records added since May 2010
LILACS	20.06.2011		10	Records published 2010 - 2011

(Continued)

WorldCat (dissertations)	20.06.2011		16	Records published 2005 - 2011
		Database total before duplicates removed	544	418
ClinicalTrials.gov	20.06.2011		14	
China Masters' Theses	15.06.2011	2000 to current	3	Searched via China National Knowledge Infrastructure Portal limited to PY 2005-2011
China Academic Journals	15.06.2011	1915 to current	19	Searched via China National Knowledge Infrastructure Portal limited to PY 2005-2011
China Doctoral Dissertations	15.06.2011	1999 to current	0	Searched via China National Knowledge Infrastructure Portal limited to PY 2005-2011
China Proceedings of Conference	15.06.2011	1999 to current	0	Searched via China National Knowledge Infrastructure Portal limited to PY 2005-2011

#### Appendix 4. Search strategies for update search in June 2011

##### Ovid MEDLINE 1948 to June Week 2 2011, searched 20 June 2011

```
1 exp Massage/
2 massag$.mp.
3 exp Touch/ or exp Therapeutic Touch/
4 touch$.mp.
5 tactile stimul$.mp.
6 or/1-5
7 (infant$ or baby or babies).mp.
8 exp Infant/
9 7 or 8
10 6 and 9
11 limit 10 to ed=20100501-20110620
```

##### CINAHL Plus (EBSCOhost) 1937 to current, searched 20 June 2011

Search limited by publication year (2005 to 2011)

```
S10 S6 and S9
S9 S7 or S8
S8 infant* or baby or babies
S7 MH Infant
S6 S1 or S2 or S3 or S4 or S5
S5 tactile stimul*
```

S4 touch\*  
S3 MH Touch or MH Therapeutic Touch  
S2 massag\*  
S1 MH Massage

**CENTRAL 2011(3), searched 20 June 2011**

#1 MeSH descriptor Massage  
#2 massag\*  
#3 MeSH descriptor Therapeutic Touch  
#4 MeSH descriptor Touch  
#5 (touch\*)  
#6 (tactile next stimul\*)  
#7 (#1 OR #2 OR #3 OR #4 OR #5 OR #6)  
#8 MeSH descriptor Infant explore all trees  
#9 infant\* or baby or babies  
#10 (#8 OR #9)  
#11 (#7 AND #10)  
#12 (#11), from 2005 to 2011  
#13 hs-handsrch  
#14 (#11 AND #13)  
#15 (#12 OR #14)

**PsycINFO (EBSCOhost) 1887 to 20 June 2011**

Search limited by publication year (2005 to 2011)  
S11 S6 and S10  
S10 S7 or S8 or S9  
S9 AG infancy  
S8 AG neonatal  
S7 infant\* or baby or babies  
S6 S1 or S2 or S3 or S4 or S5  
S5 tactile stimul\*  
S4 touch\*  
S3 DE "Tactual Perception" or DE "Tactual Stimulation"  
S2 massag\*  
S1 DE "Massage"

**EMBASE 1980 to current, searched 20 June 2011**

1. massage/  
2. massag\$.mp.  
3. exp touch/  
4. touch\$.mp  
5. tactile stimul\$.mp.  
6. or/1-5  
7. exp infant/  
8. (infant\$ or baby or babies).mp.  
9. 7 or 8  
10. 6 and 9  
11. limit 10 to yr="2005 -Current"

**LILACS searched 20 June 2011**

(Mh MASSAGE ) or (Mh TOUCH) or massag\$ or touch\$ or (tactile and stimul\$)  
 and  
 (Mh infant) or (baby or babies or infant\$)  
 and  
 (PD 2005 or PD 2006 or PD 2007 or PD 2008 or PD 2009 or PD 2010 or PD 2011)

**Maternity and Infant Care 1971 to June 2011, searched 20 June 2011. All years searched**

- 1 Massage.de.
- 2 Touch.de.
- 3 touch\$.mp.
- 4 massag\$.mp.
- 5 Therapeutic touch.de.
- 6 tactile stimul\$.mp.
- 7 or/1-6
- 8 (Infant - newborn or Infant - premature).de.
- 9 Infant - low birth weight.de. (2161)
- 10 (infant\$ or baby or babies).mp.
- 11 or/8-10
- 12 7 and 11

**WorldCat searched 20 May 2011**

(Massage or touch) AND (infants or babies)  
 Search limited by publication year (2005 to 2011) and by Content (thesis/dissertations)

**ClinicalTrials.gov searched 20 May 2011. All years searched**

(Infant\* or baby or babies) AND massage

**China Knowledge Resource Integrated Database (CNKI) searched 15 May 2011**

KW = infant or newborn AND KW =massage or therapeutic touch AND AB= random or randomly or randomised or randomised  
 Search limited by publication year 2005 to 2011

**WHAT'S NEW**

Last assessed as up-to-date: 20 December 2011.

Date	Event	Description
17 March 2013	New citation required but conclusions have not changed	Review updated with new studies and analyses
31 March 2012	New search has been performed	Updated search run. New authors



## HISTORY

Protocol first published: Issue 4, 2004

Review first published: Issue 4, 2006

Date	Event	Description
13 November 2008	Amended	Converted to new review format.
1 April 2008	Amended	Minor error about dropout in Onozawa 2001 corrected.
10 November 2006	Amended	Minor changes have been made in November 2006 (to be published Issue 1, 2007)
9 August 2006	New citation required and conclusions have changed	Substantive amendment

## CONTRIBUTIONS OF AUTHORS

This updated review was written by Jane Barlow, Cathy Bennett, Angela Underdown.

Jane Barlow will have responsibility for updating the systematic review as new material becomes available.

Cathy Bennett contributed to the updated review by reviewing search results, extracting and entering data (with other authors) and drafting the text.

## DECLARATIONS OF INTEREST

Angela Underdown - has no conflicts on interest in relation to this review.

Jane Barlow - NIHR Health Technology Assessment awarded a grant to Warwick Medical School which paid for my time in working on the update of this review.

Cathy Bennett - is the proprietor of Systematic Research Ltd and received a consultancy fee for her work on this review.

## SOURCES OF SUPPORT

### Internal sources

- University of Warwick, UK.

## External sources

- HTA, Not specified.

## DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We have rewritten sections of the [Background](#).

Consistent with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), since the previous version of the review was published, additional elements have been added to the 'Risk of bias' tables that were not present in the previous published review.

We added a comment in the [Methods](#) section, 'Unit of analysis issues' concerning cluster randomisation. None of the included studies in this review employed cluster randomisation.

A sensitivity analysis was used to assess the robustness of the findings by examining the impact of one large study (Kim 2003). This was undertaken because we were concerned that this study dominated the meta-analysis and that the results of this study may have been due to the fact that the sample comprised infants receiving orphanage care (that is, with unusually low levels of tactile stimulation), whereas the remaining studies comprised infants receiving usual levels of tactile stimulation from parents.

In this updated review, we made a further post hoc decision to record factors such as geographical location of the population and risk of bias for use in subsequent sensitivity analyses in which we repeated some of the meta-analysis, substituting alternative decisions to ensure that the results of the review are robust. We also investigated the effect of duration on intervention on outcomes and performed additional analyses accordingly.

## INDEX TERMS

### Medical Subject Headings (MeSH)

\*Child Development; \*Massage; \*Therapeutic Touch; Infant, Newborn; Physical Stimulation; Randomized Controlled Trials as Topic

### MeSH check words

Humans; Infant