SHIATSU FOR PREVENTING AND TREATING HEALTH CONDITIONS

SUPPLEMENT | ACUPRESSURE

prepared by
HTANALYSTS

for
National Health and Medical Research Council

NHMRC | Natural Therapies Working Committee
Canberra ACT 2601

JUNE 2023

Report information

Authors

Jorgensen MA[1], Taylor AE[1], Connor EM[1]

[1] **HT**ANALYSTS, Level 8, 46 Kippax Street, Surry Hills NSW 2010 Australia

Dates

This supplementary report received approval from the National Health and Medical Research Council (NHMRC) Natural Therapies Working Committee (NTWC) on 20 May 2024.

The protocol for the evidence evaluation received approval from the NHMRC NTWC on 11 March 2021 (PROSPERO: CRD42021243311).

History

NHMRC has been engaged by the Department of Health and Aged Care (the Department) to update the evidence underpinning the *2015 Review of the Australian Government Rebate on Natural Therapies for Private Health Insurance* (2015 Review) (1). The natural therapies to be reviewed are Alexander technique, aromatherapy, Bowen therapy, Buteyko, Feldenkrais, homeopathy, iridology, kinesiology, naturopathy, Pilates, reflexology, Rolfing, shiatsu, tai chi, western herbal medicine and yoga. These therapies are among those excluded from the private health insurance rebate as of 1 April 2019.

To support NHMRC in their evidence review, Health Technology Analysts (**HT**ANALYSTS) has been engaged to conduct a systematic review of the evidence of clinical effectiveness of shiatsu. Eligible studies received from the Department's public call for evidence, the Natural Therapies Review Expert Advisory Panel (NTREAP) and the NTWC will also be included in the evidence evaluation.

This supplement has been developed by **HT**ANALYSTS in conjunction with NHMRC, NTWC and NTREAP. It provides the appendices and supplementary data related to an evidence valuation of the effect of shiatsu for preventing and treating health conditions. The main body of evidence is presented in the Evidence Evaluation Report. All associated materials have been developed in a robust and transparent manner in accordance with relevant best practice standards (2-5).

Contents

List of tables

List of figures

List of abbreviations

BFI	Bowel function index
BPI	Brief pain inventory
BRISA	Regional Base of Health Technology Assessment Reports of the Americas
CINAHL	Cumulative Index to Nursing and Allied Health Literature
COMET	Core Outcome Measures in Effectiveness Trials
DBP	Diastolic blood pressure
GHQ	General health questionnaire
GRADE	Grading of Recommendations Assessment, Development and Evaluation
ITT	Intent-to-treat
MDQ	Menstrual distress questionnaire
MMSE	Mini mental state examination
NHMRC	National Health and Medical Research Council
NPRS	Numeric pain rating scale
NRSI	Nonrandomised study of an intervention
NTREAP	Natural Therapies Review Expert Advisory Panel
NTWC	Natural Therapies Working Committee
ODI	Oswestry disability index
OR	Odds ratios
PAHO	Pan American Health Organization
PICO	Population, Intervention, Comparator, Outcome
PP	Per protocol
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSQI	Pittsburgh sleep quality index
RCT	Randomised controlled trial
RMDQ	Roland-Morris disability questionnaire
RoB	Risk of bias
RR	Risk ratios
SBP	Systolic blood pressure
SR	Systematic review
SD	Standard deviation
TIDIER	Template for Intervention Description and Replication
VAS	Visual analogue scale

Supplement 1 Acupressure results

This supplement documents the systematic reviews that met the prespecified inclusion criteria for an overview[a] of the effect of acupressure for preventing and treating health conditions where results were found for a systematic review of shiatsu[b] (given acupressure's inclusion as a component of shiatsu). It provides a summary of the included reviews, a summary of the methodological quality of the reviews, and results of the data synthesis for the main comparison.

Additional details concerning the methodological quality of included systematic reviews are provided in **Appendix E2** and characteristics of the included studies are provided in **Appendix F2**. Methodological details are provided in **Appendix A** and **Appendix B**. Systematic reviews not included in the evidence synthesis are listed in **Appendix C**.

[a] A systematic review of systematic reviews

[b] See evidence evaluation report for shiatsu

SUMMARY

Acupressure is sometimes considered a central component of shiatsu and sometimes considered a therapy in its own right. As the purpose of this review is to evaluate shiatsu, evidence of the effectiveness of acupressure was included only for conditions found for shiatsu.

Detailed methods information is found in Appendix A and B. Searches were conducted in OVID, EBSCO*Host*, the Cochrane Database of Systematic Reviews, PubMed, PAHO and the Agency for Healthcare Research and Quality Systematic Review data repository. Systematic reviews provided though the Department's call for evidence were also assessed. Systematic reviews of RCTs, quasi-RCTs and NRSIs were eligible for inclusion, however only evidence from eligible RCTs (and quasi-RCTs) were included in the data synthesis. Methodological quality was assessed with AMSTAR-2 as part of the decision about which reviews to extract data. Results from RCT (and quasi-RCTs) were examined from within the systematic reviews that had been judged as providing the best available evidence. Effect estimates were reported (or calculated) from the information within the systematic reviews, with combined results included where available. Information about RCTS was checked across reviews where possible. Evidence was assessed using the Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) framework (5). Summary of Findings tables were developed for studies that compared acupressure to either sham (comparison 1) or inactive control (comparison 2) and which reported on outcomes rated as critical or important by NTWC. The summary of Findings tables included information about primary studies provided within a systematic review (e.g., study design, population characteristics, risk of bias) supplemented with additional information provided across reviews or developed within the overview (e.g., publication bias).

For the supplementary overview of acupressure, 93 systematic reviews were identified as eligible for inclusion (i.e. 66 relevant SRs covering conditions that were also identified in the systematic review for the effectiveness of shiatsu and 27 partially relevant SRs[c] covering these conditions). At the time of the search, an additional 25 reviews were awaiting classification, and an additional 11 reviews were recorded as ongoing (protocol registered but not complete). Of the studies awaiting classification, 13 were published in languages other than English, 11 were conference abstracts or posters and 1 review was not able to be retrieved.

Within these SRs, 90 RCTs covering 9 conditions were considered in the evidence evaluation and are included in the results. For the synthesis, there were 49 RCTs covering 8 conditions that compared acupressure with a sham intervention and 41 RCTs covering 7 conditions that compared acupressure with an inactive control (no intervention, waitlist or usual care).

More than one-third of the primary studies included in the synthesis were in pregnant females (34 studies), with the remaining primary studies identified covering 8 other conditions (between 2 and 14 primary studies per condition). There were no primary studies found within the eligible systematic reviews for 7 conditions that had been included in the systematic review for shiatsu, including diabetes, obesity, stress, headache disorders, stroke recovery, postpartum care, and burn injuries.

For the supplementary overview of acupressure all forms applied to traditional Chinese medicine acupuncture (tsubo) points were eligible for inclusion. The acupressure could be either individualised or non-individualised and could be self-administered or applied by a therapist or lay person to specific points using a finger, hand, elbow, or foot. The acupressure could also be delivered via an acupressure band or large bead as a replacement for finger pressure.

All included studies examined acupressure exercises delivered in a manner that was applicable to the Australian context based on the description. Most studies evaluated acupressure that was self-administered to specified acupoints. In other cases, acupressure was delivered via acupressure bands or via an experienced therapist or nurse.

This overview identified 9 conditions for which there was evidence about the effect of acupressure on an outcome considered critical or important by NTWC.

[c] The SR included a mix of both relevant and irrelevant primary studies.

Compared with a sham intervention, the evidence provides:

- moderate certainty that acupressure probably results in:
  - a large reduction in post-operative vomiting after minimally invasive surgery (8 RCTs, 597 participants)
  - a slight improvement in sleep quality in people with insomnia (4 RCTs, 213 participants)
- low certainty that acupressure may result in:
  - a large reduction in labour duration (6 RCTs, 559 participants) and a reduction in labour pain (9 RCTs, 935 participants) in pregnant females
  - a reduction in anxiety (2 RCTs, 175 participants) in cancer survivors
  - an improvement in quality of life in people with insomnia (1 RCT, 62 participants)
  - a reduction in pain intensity in females with dysmenorrhoea (4 RCTs, 380 participants)
  - a slight reduction in post-operative nausea after minimally invasive surgery (8 RCTs, 606 participants)
- low certainty that acupressure provides little (to no) benefit in:
  - improving psychosocial wellbeing in people with insomnia (1 RCTs, 40 participants).

The evidence provides very low certainty of the effect of acupressure compared with sham for 10 out of the 100 critical or important outcomes prioritised for analysis in this review[d]. Of the 100 outcomes prioritised as critical or important in this review, 81 were not addressed by any studies, and therefore the effect of acupressure on these 81 outcomes compared with a sham intervention is unknown.

Compared with an inactive control (no intervention, waitlist or usual care), the evidence provides:

- moderate certainty that acupressure is effective in:
  - providing a slight improvement in sleep quality in people with insomnia (2 RCTs, 125 participants)
- low certainty that acupressure provides:
  - a large improvement in fatigue in cancer survivors (1 RCT, 158 participants)
  - an improvement in symptom severity in people with functional constipation (1 RCT, 100 participants)
  - an improvement in pain intensity in people with dysmenorrhoea (5 RCTs, 363 participants)
  - in reduction in labour duration (4 RCTs, 338 participants) and a reduction in labour pain (8 RCTs, 615 participants)
  - a slight improvement in neurocognitive function in people with neurocognitive disorders (1 RCT, 76 participants)
- low certainty that acupressure provides little (to no) benefit in:
  - reducing nausea in people with cancer (2 RCTs, 144 participants)
  - improving quality of life in people with functional constipation (1 RCT, 100 participants)
  - reducing symptom severity in people with dysmenorrhoea (3 RCTs, 380 participants).

Compared with an inactive control, the evidence provides very low certainty of the effect of acupressure for 8 out of the 100 critical or important outcomes prioritised for analysis in this review. For these outcomes, the true effect is probably markedly different from the estimated effect, with more studies needed to determine the true effect. Of the 100 outcomes prioritised as critical or important in this review, 82 were not addressed by any studies, and therefore the effect of acupressure on these 82 outcomes compared with an inactive control is unknown.

A summary of harms of acupressure is not possible, as it was out of scope of this review to assess adverse effects related to acupressure.

[d] Across 5 conditions (cancer survivors, neurocognitive decline, hypertensive heart disease, chronic musculoskeletal pain, dysmenorrhoea)

Overall, acupressure may provide people who receive it with some benefit for a small number of relevant outcomes (up to three for a given condition) when compared with sham or an inactive control. In some cases, the true size of the effect estimate was uncertain (5 outcomes) or unknown (74 outcomes). Apart from 2 conditions (pregnancy, recovery after minimally invasive surgery) almost all the effect estimates were based on results from fewer than 4 RCTs (range 22 to 380 total participants) which can impact the precision of the results. For a few outcomes, a clinically important difference was not observed (possibly relating to study design, size or duration).

In considering the use of acupressure as part of the overall practice of shiatsu, it is difficult to provide guidance. Compared with an inactive control (no intervention) the results for acupressure were often inconsistent with the effect reported for shiatsu or a clear judgement about consistency of the effect could not be made as the certainty of evidence for acupressure (or shiatsu) was very low or unknown. Compared with a sham intervention, the effect of shiatsu is unknown – with results of this overview indicating acupressure improves some outcomes for some conditions and not others. Overall, these conclusions are sometimes based on a small number of studies with limited numbers of participants, with results across studies often imprecise and inconsistent.

S1.1 Description of studies

S1.1.1 Flow of studies

The literature was searched on 21 April 2021 to identify relevant studies published from database inception to the literature search date. The results of the search and application of the study selection criteria are provided in Appendix A3.2 – A5 and Appendix C1 and C2.

A PRISMA flow diagram summarising the screening results is provided in Figure S1. The flow diagram shows the number of studies at each stage of search and screening process, including: the initial search; studies considered irrelevant based on the title and/or abstract; studies found not to be relevant when reviewed at full text; studies which met the eligibility criteria for inclusion in the review and the number of studies which were in considered in the analysis for conditions matching those found for shiatsu.

The search retrieved 66 systematic reviews that covered populations identified in the systematic review for the effectiveness of shiatsu and were eligible for inclusion (see Included studies). No additional reviews were identified and included from the Department's public call for evidence. A further 25 reviews are awaiting classification and 11 reviews were recorded as ongoing.

S1.1.2 Excluded studies

There were 104 citations screened at full text that were excluded for not meeting the eligibility criteria of this review. Of these, 57 had an intervention out of scope (e.g. not acupressure or unable to assess acupressure independent of other interventions), 15 had a study design out of scope (i.e. systematic review of systematic reviews), 15 had been superseded (i.e. a newer version of the systematic review was available), 8 had a publication type out of scope (e.g. opinion piece or not an interventional study), 4 had been withdrawn (i.e. the systematic review no longer met Cochrane standards or expectations), 2 were in a population out of scope (i.e. healthy population not at risk), one examined outcomes that were out of scope (patient experience), and one was a duplicate report of the same published data.

Details of citations which were thought likely to be eligible but were not, are presented in Appendix C1.2. Some studies may have been out of scope for more than one reason, but only one reason is listed for each.

S1.1.3 Studies awaiting classification

Completed studies identified as potentially eligible for inclusion that could not be retrieved, translated or provided insufficient or inadequate data, are listed in the *Characteristics of studies awaiting classification* tables (see Appendix C3.2). This includes 11 conference proceedings with incomplete information about the study (Appendix C3.2.1), 13 reviews published in languages other than English (Appendix C3.2.2) that are possibly eligible for inclusion (pending translation into English), and one review that was not able to be retrieved (Appendix C3.2.3).

The 25 studies awaiting classification were comparable to those included in the evidence synthesis in terms of conditions examined and outcomes measured.

S1.1.4 Ongoing studies

There were 11 ongoing reviews that did not have published results at the time of the search. These are listed in the *Characteristics of ongoing studies* table (see Appendix C4.2).

S1.1.5 Included studies

An overview of the conditions identified and included in this review is provided in Table S1.

There were 126 systematic reviews identified as eligible for inclusion in the review. Of these, 66 systematic reviews covered conditions that were also identified in the systematic review for the effectiveness of shiatsu (see main report) and were considered in the evidence synthesis for acupressure. Detailed descriptions of the included reviews, including a summary of the PICO criteria of included reviews, critical appraisal, applicability of the studies to shiatsu and results of the data synthesis for the main comparisons is provided below. An inventory of systematic reviews that covered conditions not identified in the evidence review for shiatsu is provided in Appendix C5.

For comparison 1 (acupressure compared with sham), 49 RCTs from within the systematic reviews were considered in the synthesis. For comparison 2 (acupressure compared with no intervention, waitlist or usual care, if considered inactive), 41 RCTs from within the systematic reviews were considered for synthesis. Primary studies that included NTWC prioritised critical and important outcome domains and measures (highlighted in a blue box in Appendix F1.2), were included in the final analysis. Details about the RCTs that compared shiatsu with other (active) comparators are included in qualitative descriptions in the report, but results from these studies were not extracted.

Figure S1 Literature screening results: Acupressure

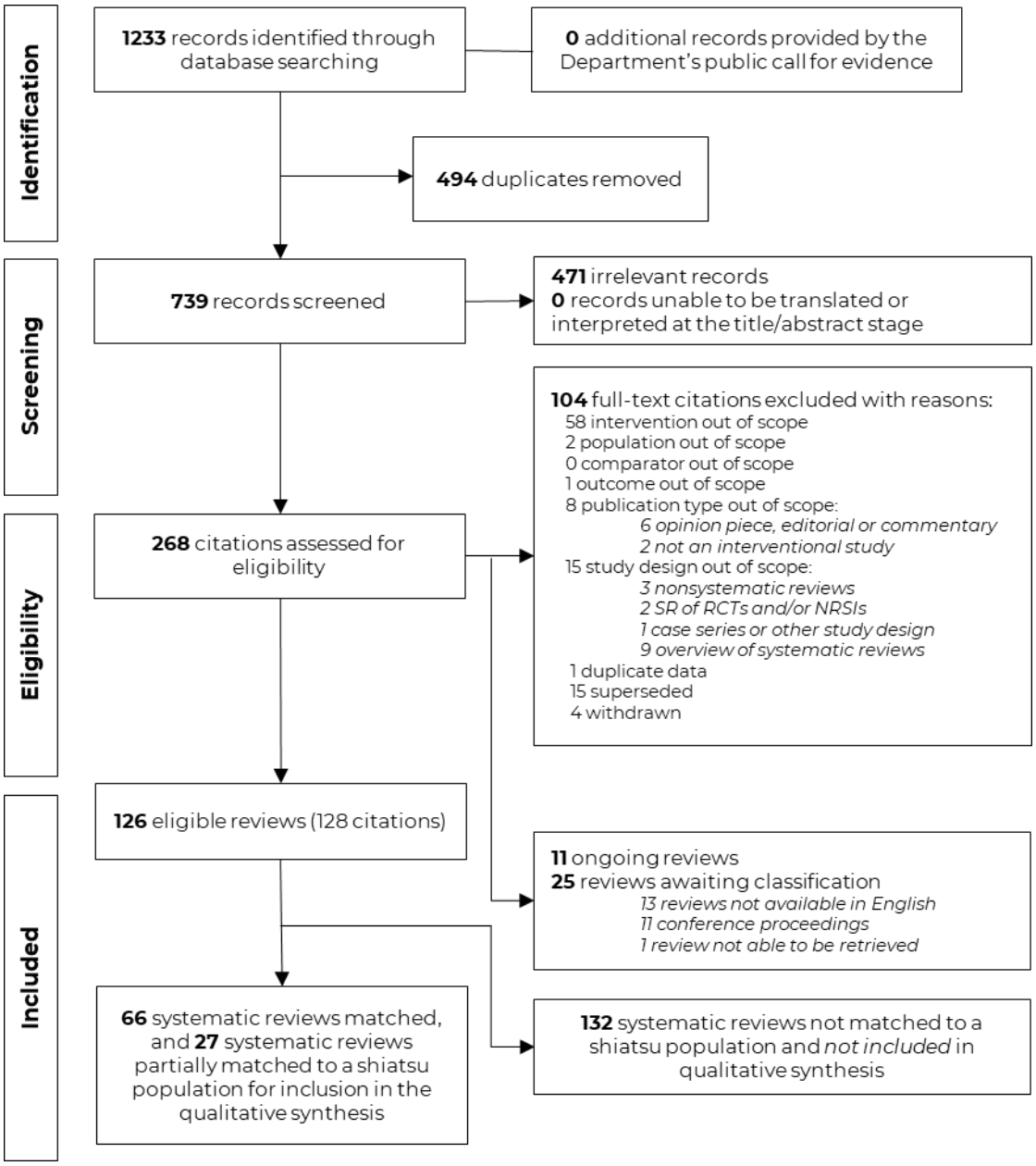


Table S1 List of conditions and population groups in identified systematic reviews

ICD-11 [a]	POPULATION	Number of reviews [b]	MATCHED TO CONDITION IDENTIFIED IN SHIATSU		
			No	Partial	Yes
02 Neoplasms [c]					
	Cancer, on treatment (any type, breast, stomach, lung, mixed)	18	18		
	Cancer, survivors (any type)	0			
	Cancer, not specified (any type, including after haematopoietic stem cell transplant)	6		6	
05 Endocrine, nutritional and metabolic diseases					
	Diabetes, type 2	1			**1**
	Hyperthyroidism	1	1		
06 Mental and behavioural disorders					
	21 Symptoms of anxiety	4	4		
	Mood disorders	4	4		
	Neurocognitive, Alzheimer's disease, dementia and/or mild cognitive impairment	9		8	**1**
	Neurodevelopmental, autism spectrum disorders	2	2		
	Nocturnal enuresis (children)	2	2		
	Substance abuse, nicotine	1	1		
07 Sleep-wake disorders					
	21 Sleep disturbance	13			**13**
	Insomnia	2			**2**
08 Diseases of the nervous system					
	21 Vertigo (acute)	1	1		
	Epilepsy	1	1		
	Headache disorders	3			**3**
	Neurological disorders, mixed (Parkinson's disease, multiple sclerosis, stroke recovery)	1	1		
	Postviral fatigue syndrome	1	1		
	Stroke recovery	4			**4**
11 Diseases of the circulatory system					
	Hypertensive heart disease	5			**5**
	Ischaemic heart disease (angina, peripheral arterial occlusive disease, coronary heart disease, acute myocardial infarction)	5	5		
12 Diseases of the respiratory system					
	Asthma	3	3		
	Bronchiectasis	4	4		
	COPD	7	7		
	Rhinitis, allergic	2	2		
13 Diseases of the digestive system					
	Functional constipation	2			**2**
14 Diseases of the skin					
	Dermatitis, atopic	3	3		
	Pruritis, uraemic (due to end-stage renal disease)	3	3		

ICD-11 [a]	POPULATION	Number of reviews [b]	MATCHED TO CONDITION IDENTIFIED IN SHIATSU		
			No	Partial	Yes
15 Diseases of the musculoskeletal system or connective tissue					
	21 Low back pain	7			**7**
	21 Musculoskeletal pain	2			**2**
	Osteoarthritis, knee	3	3		
	Spondylosis, cervical	1	1		
16 Diseases of the genitourinary system					
	Chronic kidney disease	16	16		
	Dysmenorrhoea and/or menstrual distress	18			**18**
	Premenstrual syndrome or premenstrual dysphoric disorder	1	1		
	Symptoms of menopause	2	2		
	Urinary incontinence (stress)	1	1		
18 Pregnancy, childbirth or the puerperium [d]					
	Active labour (pain management)	10	9	1	
	Breastfeeding mothers (mastitis, lactation)	2	2		
	Labour induction / duration and mode	6			**6**
	Pregnant women (nausea and vomiting)	8	8		
19 Certain conditions originating in the perinatal period					
	Neonatal jaundice	1	1		
22 Injury, poisoning or certain other consequences of external causes					
	Acute injury (ankle sprain, minor trauma, contusions or fracture)	7	7		
	Fracture, hip	1	1		
	Traumatic Brain Injury	1	1		
24 Factors influencing health status or contact with health services					
	Acute pain (associated with intramuscular injection)	1	1		
	Caregivers, family and/or informal (sleep)	1	1		
	Palliative care	1	1		
	Recovery after surgery[e] (applied after surgery) (nausea & vomiting, sleep, pain, gastrointestinal motility) including: - eye surgery (children) - minimally invasive abdominal - endoscopic retrograde cholangiopancreatography - knee arthroscopy - Caesarean section - coronary artery bypass surgery - transabdominal hysterectomy	18	4	12	**2**
25 Prevention					
	Gag reflex, patients undergoing dental treatment	1	1		
	Preoperative anxiety	1	1		
	Post-operative nausea and vomiting (applied before anaesthesia during breast surgery or Caesarean section or simulated motion sickness)	7	7		
Grand Total		**225**	**132**	**27**	**66**

a. International Statistical Classification of Diseases and Related Health Problems 11th Revision (ICD-11)-WHO Version (2021)
b. Numbers reflect the population considered *within* the systematic review and not the number of included systematic reviews (i.e. umbrella reviews that considered more than one population are counted more than once).

c. Systematic reviews that focused on acupressure in people receiving chemotherapy, in palliative care, or mixed populations were not included as priority as the evidence in shiatsu was focused on cancer survivors.
d. Systematic reviews that focused on acupressure in early pregnancy (nausea & vomiting) or only on labour pain were not prioritised as the evidence in shiatsu was focused on labour induction.
e. Systematic reviews that focused on acupressure in people recovering after major surgery (e.g. cardiopulmonary bypass, C-section) were not included here as evidence in shiatsu was focused on minimally invasive procedures.

S1.2 Cancer (survivors)

S1.2.1 Description of studies

Seven citations (6-12) corresponding to 7 systematic reviews (Lee 2011c, Ling 2014, Duong 2017, Arring 2019, Calcagni 2019, Harvie 2019, Liu 2020) were identified in the literature that assessed acupressure compared to sham, control or an active intervention in people with cancer[e]. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are no systematic reviews awaiting classification (see Appendix C3.2) and one ongoing review (13) (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S2. Review details, including all outcome domains and measures reported by the SR, and the risk of bias of the included primary studies are provided in Appendix F1.2.

The primary studies included by the systematic review authors had been conducted in people with a variety of cancer diagnoses (e.g. lung, breast, undergoing bone marrow biopsy) and it was not always clear when acupressure treatment was applied (i.e. before, during or after treatment). For these reasons, the available evidence may not be directly applicable to the population considered in the shiatsu evidence review (gynaecological cancer survivors), but it could be sensibly applied.

There were 7 studies (Molassiotis 2007, Tang 2014, Beikmoradi 2015, Avci 2016, Rizi 2017, Zhang 2017, Hoang 2019) that assessed acupressure compared to a sham intervention, and 7 studies (Beikmoradi 2015, Hsiung 2015, Hughes 2015, Zick 2016, Nia 2017, Rizi 2017, Hoang 2019) that compared acupressure with control (no intervention or usual care). One study (Hughes 2015) also included an active intervention group (auricular acupressure).

Table S2 List of included systematic reviews and overlap with eligible RCTs (per outcome): Cancer

Review ID	Best available*	SR Outcome domain (measure)	Study ID										
			Molassiotis 2007	Tang 2014	Beikmoradi 2015	Hsiung 2015	Hughes 2015	Avci 2016	Zick 2016	Nia 2017	Rizi 2017	Zhang 2017	Hoang 2019
Lee 2011c (6)	†	Fatigue (BFI)	?	--	--	--	--	--	--	--	--	--	--
Ling 2014 (7)	†	Fatigue (BFI or other)	Y	--	--	--	--	--	--	--	--	--	--
Duong 2017 (8)	✓	Fatigue (self-reported) **	Y	Y	--	--	--	--	Y	--	--	--	--
Arring 2019 (9)	X	Fatigue (BFI)	--	--	--	--	--	--	Y	--	--	--	--
Calcagni 2019 (10)	†	Fatigue (NR)	--	--	!	!	--	!	--	!	!	?	--
		Mood (NR)	--	--	!	!	--	!	--	!	!	?	--
		Nausea (NR)	--	--	!	?	--	?	--	!	!	!	--
		Psychosocial wellbeing (NR)	--	--	Y	!	--	!	--	!	Y	!	--
		Pain (NR)	--	--	!	?	--	!	--	?	Y	!	--
		Sleep disturbance (NR)	--	--	!	!	--	!	--	!	!	?	--
Harvie 2019 (11)	†	Fatigue (BFI)	--	--	--	--	!	--	Y	--	--	--	--
		Sleep quality (PSQI)	--	--	--	--	Y	--	Y	--	--	--	--
Liu 2020 (12)	†	Sleep disturbance (PSQI)	--	Y	--	--	--	--	Y	--	--	--	Y

Abbreviations: BFI, brief fatigue inventory; ID, identification; NR, measure not reported; PSQI, Pittsburgh sleep quality index

* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 9, domain 8 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])

[e] Systematic reviews that focused on acupressure in people receiving chemotherapy, in palliative care, or mixed populations were not included as priority as the evidence in shiatsu was focused on cancer survivors (see Appendix C5).

✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
** A range of measures were used including the brief fatigue inventory, Piper Fatigue Scale and VAS. The review did not specify which study used what measure.
Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]
? RCT is included in the systematic review, meets our PICO criteria but a study result is not available for the listed outcome measure [data is incomplete; result may be available in another SR]
! RCT is included in the systematic review but the SR indicates that study does not report the listed outcome [not measured]
-- RCT is not included in the systematic review

S1.2.2 Critical appraisal

Out of 6 included systematic reviews, one review (Duong 2017) was judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest (i.e. met, or partially met, critical AMSTAR-2 domains 4, 8, 9 and 11).

The other 5 reviews (Arring 2019, Calcagni 2019, Harvie 2019, Liu 2020, Ling 2014) had at least one critical flaw (i.e. did not meet, or partially meet, one of the prespecified critical AMSTAR-2 domains [4, 8, 9 or 11]). Of these, 3 reviews (Arring 2019, Ling 2014, Liu 2020) did not use a comprehensive literature search (domain 4), one review (Arring 2019) did not appropriately assess risk of bias (domain 9), and the other systematic reviews did not conduct a meta-analyses (i.e. provided a narrative review of individual studies).

A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.2.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in cancer survivors are listed in Table S3.

Table S3 Outcomes considered by the NTWC to be critical or important for decision-making: Cancer (survivors)

Prioritised outcome domain	Measured with	Consensus rating	Results available for comparison 1 or 2?	Review ID					
				Ling 2014	Duong 2017	Arring 2019	Calcagni 2019	Harvie 2019	Liu 2020
Quality of life	EORTC QLQ-C30 (total)	Critical	No	?	?	?	--	--	--
Pain	Visual analogue scale	Critical	Yes	?	?	?	X	--	?
Physical symptoms	EORTC QLQ-C30 (nausea & vomiting)	Critical	Yes	?	?	?	†	--	?
Fatigue	EORTC QLQ-C30 (fatigue)	Critical	Yes	*	✓	*	*	*	?
Physical functioning	EORTC QLQ-C30 (physical functioning)	Critical	No	?	?	?	--	--	?
Overall wellbeing	EORTC QLQ-C30 (overall wellbeing)	Critical	No	?	?	?	--	--	--
Psychosocial wellbeing	Hospital Anxiety and Depression Scale	Important	Yes	?	?	?	†	--	?

Abbreviations: EORTC QLQ-C30, 30-item European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire
✓ A study result is available for inclusion in the synthesis
* A study result is available and is reported by another systematic review nominated as the best available evidence.
X A study result is available for inclusion, but the systematic review reported incomplete data. Due to time and resource constraints, only the information presented in the systematic review is reported.
† A study result is available for inclusion, but the systematic review only reports the direction of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.
-- The systematic review assessed this outcome but did not find or include any eligible primary studies reporting the outcome.

? The systematic review did not assess this outcome. It is unclear if the outcome was assessed by the primary studies included in the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

Two systematic reviews (Calcagni 2019, Duong 2017) included evidence from 3 RCTs comparing acupressure with sham in cancer (survivors) that contributed data to one of the 7 critical or important outcomes. Calcagni 2019 also reports data from one RCT, in which the sham and control groups were combined[f] (method or reasons not described).

Pain

One primary study (Rizi 2017) was identified by Calcagni 2019 that reported pain (total 90 participants). The measure used, and treatment duration were not reported. The systematic review authors suggest an effect in favour of acupressure compared with the combined sham or control (SMD –1.6; 95% CI –2.2, –1.0; *p* = not reported).

Fatigue

There were 2 primary studies (Molassiotis 2007, Tang 2014) identified by one systematic review (Duong 2017) that reported fatigue (total 65 participants) measured with either the brief fatigue inventory, Piper Fatigue Scale or a visual analogue scale (VAS) at the end of treatment (range 2 to 20 weeks). The specific measure used in each study was not reported. Pooled results suggest an effect in favour of acupressure compared with the sham group (SMD –0.96; 95% CI –1.88, –0.03; *p* = 0.04).

One additional primary study (Zhang 2017) was identified by another systematic review (Calcagni 2019) that measured fatigue (total 43 participants). The measure used and treatment duration were not reported. The review authors reported that data from the study were not available, but that the results suggested an effect in favour of acupressure.

Psychosocial wellbeing (anxiety)

There were 2 primary studies (Beikmoradi 2015, Rizi 2017) identified by Calcagni 2019 that reported psychosocial wellbeing (anxiety) (total 175 participants). The measure used and treatment duration were not reported.

The systematic review authors report individual study results comparing acupressure with sham/control at the end of treatment, both of which suggest an effect in favour of acupressure (Beikmoradi 2015: SMD –0.9; 95% CI –1.3, –0.4; Rizi 2017: SMD –0.5; 95% CI –1.1, –0.04; *p* = not reported).

Comparison 2 (vs control)

Two systematic reviews (Calcagni 2019, Duong 2017) included evidence from 4 primary studies comparing acupressure with control (no intervention, usual care) in cancer survivors that contributed data to 3 of the 7 critical or important outcomes.

Pain

There were 2 primary studies (Hsiung 2015, Nia 2017) identified by one systematic review (Calcagni 2019) that reported pain (total 154 participants), but the measures used and treatment duration were not reported. The review authors reported that data from the studies were not available but that the results showed a positive effect in favour of acupressure in one study (Hsiung 2015) and the results suggested no difference between groups in one study (Zhang 2017).

Physical symptoms (nausea and vomiting)

There were 2 primary studies (Hsiung 2015, Avci 2016) identified by one systematic review (Calcagni 2019) that reported nausea (total 144 participants), but the measures used and treatment duration were not reported. The review authors reported that data from the studies were not available but that the results showed no difference between treatment groups.

[f] The protocol did not describe how combined data would be included in the evidence synthesis. We have elected to include with Comparison 1.

Fatigue

One primary study (Zick 2016) was identified by one systematic review (Duong 2017) that reported fatigue (total 158 participants) measured with either the brief fatigue inventory, Piper Fatigue Scale or VAS at the end of treatment (3 weeks). The specific measure used was not reported. The results suggested an effect in favour of acupressure compared with the control group (SMD –0.82; 95% CI –1.15, –0.50; $p < 0.00001$).

Comparison 3 (vs other)

There were no systematic reviews found that included RCTs comparing acupressure with an active comparator in cancer survivors that provided any data for the critical or important outcomes.

S1.2.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There were 3 RCTs found by the included systematic reviews comparing acupressure with sham in people with cancer that contributed data to one prioritised outcome (fatigue). One other RCT contributed data to 2 other prioritised outcomes (pain, anxiety) (in which the results for the sham and control groups had been combined).

Acupressure compared to sham for Cancer (survivors)

Patient or population: Cancer (survivors)
Setting: community
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on health-related quality of life in people with cancer is unknown
Pain assessed with: not reported Follow-up: not reported	-	SMD **1.6 SD lower^** (2.2 lower to 1.0 lower)		90 (1 RCT) **	⊕◯◯◯ VERY LOW a,b,c,d,e	The evidence is very uncertain about the effect of acupressure on pain in people with cancer
Physical symptoms – not reported	-	-	-	(0 studies)	-	The effect of acupressure on physical symptoms in people with cancer is unknown
Physical functioning – not reported	-	-	-	(0 studies)	-	The effect of acupressure on physical functioning in people with cancer is unknown
Fatigue assessed with: BFI, VAS or other Follow-up: range 2 to 20 weeks	-	SMD **0.96 SD lower^** (1.88 lower to 0.03 lower)	-	65 (2 RCTs) # missing data from 1 RCT	⊕◯◯◯ VERY LOW a,c,e,f,g	The evidence is very uncertain about the effect of acupressure on fatigue in people with cancer
Overall wellbeing – not reported	-	-	-	(0 studies)	-	The effect of acupressure on overall wellbeing in people with cancer is unknown

Acupressure compared to sham for Cancer (survivors)

Patient or population: Cancer (survivors)
Setting: community
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Psychosocial wellbeing (anxiety) assessed with: not reported Follow-up: not reported	Individual study results: SMD **0.9 SD lower^** (1.13 lower to 0.4 lower) SMD **0.5 SD lower^** (1.1 lower to 0.04 lower)			175 (2 RCTs) **	⨁⨁◯◯ LOW [a,c,e,f,h]	Acupressure may result in a reduction in anxiety in people with cancer
Psychosocial wellbeing (depression) – not reported	-	-	-	(0 studies)	-	The effect of acupressure on depression in people with cancer is unknown

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

** evidence is from RCTs with sham and control group results combined.
^ As a rule of thumb, an SMD of 0.2 is considered a small difference, 0.5 is medium, and 0.8 is large difference. (14).
# Data from one RCT (45 participants) not included. The review authors report the primary study did not provide complete data, but that an effect in favour of acupressure was observed.

BFI: Brief Fatigue Inventory; **CI:** confidence interval; **MD:** mean difference; **VAS:** visual analogue scale

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. Single study. Inconsistency not assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with various cancers (both on and off treatment) and is probably generalisable to cancer survivors with some caveats. Certainty of evidence not downgraded.
d. Very serious imprecision. Single study with combined sham and control groups. Wide confidence intervals. Certainty of evidence downgraded 2 levels.
e. Publication bias suspected. Evidence is limited to a small number of small trials. Certainty of evidence downgraded.
f. No serious inconsistency. Certainty of evidence not downgraded.
g. Very serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and trivial or no important difference). Certainty of evidence downgraded 2 levels.
h. Serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and small [or trivial] important difference). Certainty of evidence downgraded.

Comparison 2 (vs control)

There were 4 RCTs found by the included systematic reviews comparing acupressure with control (no intervention, waitlist, usual care) in people with cancer that contributed data to 3 prioritised outcomes.

Acupressure compared to control (no intervention, waitlist, usual care) for Cancer (survivors)

Patient or population: Cancer (survivors)
Setting: community
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on health-related quality of life in people with cancer is unknown
Pain assessed with: not reported Follow-up: not reported	Data not available for 2 RCTs that report conflicting results.		-	154 (2 RCTs)	⊕◯◯◯ VERY LOW [a,c,d,e,f]	The evidence is very uncertain about the effect of acupressure on pain in people with cancer
Nausea assessed with: not reported Follow-up: not reported	Data not available for 2 RCTs that report no difference between treatment groups		-	144 (2 RCTs)	⊕⊕◯◯ LOW [a,b,c,d,e]	Acupressure may result in little to no difference on nausea in people with cancer
Fatigue assessed with: not reported Follow-up: not reported	-	SMD **0.82 SD lower^** (1.15 lower to 0.50 lower)	-	158 (1 RCT)	⊕⊕◯◯ LOW [a,b,c,d,e]	Acupressure may result in a large reduction in fatigue in people with cancer
Physical functioning – not reported	-	-	-	(0 studies)	-	The effect of acupressure on physical functioning in people with cancer is unknown
Overall wellbeing – not reported	-	-	-	(0 studies)	-	The effect of acupressure on overall wellbeing in people with cancer is unknown
Psychosocial wellbeing (stress, anxiety or depression) – not reported	-	-	-	(0 studies)	-	The effect of acupressure on psychosocial wellbeing in people with cancer is unknown

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

^ As a rule of thumb, an SMD of 0.2 is considered a small difference, 0.5 is medium, and 0.8 is large difference (14).

CI: confidence interval; **MD:** mean difference; **PAC-QoL:** Patient Assessment of Constipation – Quality of life

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

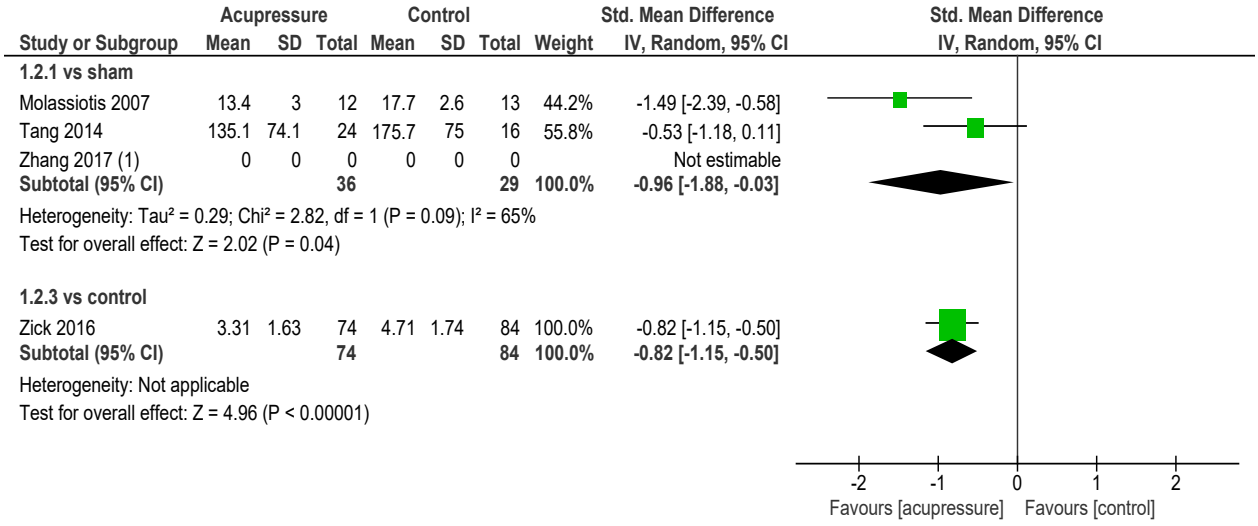
Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency or inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with various cancers (both on and off treatment) and is probably generalisable to cancer survivors with few caveats. Certainty of evidence not downgraded.
d. Serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and no important difference). Certainty of evidence downgraded.
e. Publication bias suspected. Evidence is limited to a small number of small trials. Certainty of evidence downgraded.
f. Serious inconsistency. One study suggests an effect in favour of acupressure and one study suggests no difference between groups. Certainty of evidence downgraded.

S1.2.5 Forest plots

Outcome results for people with cancer (where additional analyses were required and able to be carried out) are presented in Figure S2 (fatigue).

Figure S2 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): Cancer – fatigue



Footnotes
(1) Data not available (total 43 participants). Review authors suggest effect in favour of acupressure.

S1.3 Diabetes

S1.3.1 Description of studies

One citation (15) corresponding to one systematic review (Robinson 2011) was identified in the literature that assessed acupressure compared to sham, control or an active intervention in people with diabetes. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are no systematic reviews awaiting classification (see Appendix C3.2) and no ongoing reviews (see Appendix C4.2).

Review details, including all outcome domains and measures, and the risk of bias of the included studies are provided in Appendix F1.2.

The review authors identified one RCT (Jin 2009) that examined the effect of acupressure for treatment of symptoms associated with diabetes. The RCT was reported to be of high quality with results suggesting improvements in hyperlipidaemia, ventricular hypertrophy, kidney function and neuropathy but no other information was provided.

In the absence of any usable data, the review/study were not further considered.

S1.3.2 Critical appraisal

A summary of the strengths or limitations of the included systematic review assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.4 Obesity

S1.4.1 Description of studies

One citation (16) corresponding to one systematic review (Ernst 1997) was identified in the literature that assessed acupressure compared to sham, control or an active intervention in people with obesity (or overweight). The review was published in a language other than English and is therefore awaiting classification (see Appendix C3.2). No additional reviews were identified in the Departments public call for evidence (see Appendix C2) and there are no ongoing reviews (see Appendix C4.2).

S1.5 Neurocognitive decline

S1.5.1 Description of studies

There were 8 citations (15, 17-23) corresponding to 8 systematic reviews (Lee 2011a, Robinson 2011, Strom 2016, Liu 2018, Hmwe 2019, Margenfield 2019, O'Caoimh 2019, Chen 2020a) identified in the literature that assessed acupressure compared to sham, control or an active intervention in people with neurocognitive disorders. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are no systematic reviews awaiting classification (see Appendix C3.2) and no ongoing reviews (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S4. Review details, including all outcome domains and measures, and the risk of bias of the included studies are provided in Appendix F1.2.

The studies included by the systematic review authors were conducted in people with dementia or mild cognitive disorders and may not be directly applicable to the population considered in the shiatsu evidence review (Alzheimer's) but could be sensibly applied. Two studies (Kwan 2017, Mariko 2015) compared acupressure (or acupoint massage) with sham, and 4 studies (Feng 2015, Lin 2009, Sun 2016, Yang 2007) compared acupressure with control (no intervention, waitlist, usual care). The comparator details for one study (Wan 2017) were not provided. There were 2 studies that also included an active intervention group (Kwan 2017, Lin 2009).

Table S4 List of included systematic reviews and overlap with eligible RCTs (per outcome): Neurocognitive decline

Review ID	Best available*	SR Outcome domain (measure)	Study ID						
			Yang 2007	Lin 2009	Feng 2015	Mariko 2015	Sun 2016	Kwan 2017	Wan 2017
Lee 2011a (17)	†	Behavioural symptoms (CMAI)	?	--	--	--	--	--	--
Robinson 2011 (15)	†	Behavioural symptoms (CMAI)	?	?	--	--	--	--	--
Strom 2016 (18)	†	Behavioural symptoms (CMAI)	--	?	--	--	--	--	--
Liu 2018 (19)	✓	Neurocognitive function (MMSE)	--	--	?	--	Y	--	--
Hmwe 2019 (20)	✓	Behavioural symptoms (CMAI)	--	?	--	--	--	?	--
		Biomarkers (salivary cortisol)	--	!	--	--	--	?	--
Margenfield 2019 (21)	✓	Behavioural symptoms (CMAI)	--	--	--	!	--	Y	--
		Biomarkers (salivary cortisol)	--	--	--	!	--	Y	--
		Behavioural symptoms (NPI)	--	--	--	Y	--	!	--
		Functional capability (ADL)	--	--	--	Y	--	!	--
		Neurocognitive function (MMSE)	--	--	--	Y	--	!	--
O'Caoimh 2019 (22)	†	Sleep quality	No eligible studies found						
Chen 2020a (23)	†	Neurocognitive function (MMSE)	--	--	--	--	Y	--	Y

Abbreviations: ADL, activities of daily living; CMAI, Cohen Mansfield agitation inventory; GDS, geriatric depression scale; NPI, neuropsychiatric inventory; MMSE, mini mental state examination; PSQI, Pittsburgh sleep quality index

* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 8, domain 9 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])

✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]

? RCT is included in the systematic review, meets our PICO criteria but a study result is not available for the listed outcome measure [data is incomplete; result may be available in another SR]

! RCT is included in the systematic review, but the SR indicates that the study does not measure the listed outcome [not measured]
-- RCT is not included in the systematic review

S1.5.2 Critical appraisal

Out of 7 systematic reviews, 3 reviews (Hmwe 2019, Liu 2018, Margenfield 2019) were judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest (i.e. met, or partially met, the prespecified critical AMSTAR-2 domains [4,8,9, and 11]).

The other 5 systematic reviews (Lee 2011a, Robinson 2011, Storm 2016, O'Caoimh 2019, Chen 2020a) had at least one critical flaw (i.e. did not meet, or partially met, one of the prespecified critical AMSTAR-2 domains [4, 8, 9 or 11]). Of these, 4 systematic reviews (Lee 2011a, Storm 2016, O'Caoimh 2019) did not conduct a comprehensive literature search (domain 4), 4 reviews (Lee 2011a, Robinson 2011, O'Caoimh 2018, Storm 2016) did not use a satisfactory technique for assessing the risk of bias of individual studies (domain 9), two systematic reviews (Robinson 2011, Chen 2020a) failed to adequately describe the included studies in detail (domain 8) and 4 systematic reviews (Lee 2011a, Robinson 2011, Hmwe 2019, Storm 2016) provided narrative summaries only (no meta-analysis) (domain 11).

A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.5.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in people with neurocognitive disorders are listed in Table S5.

Table S5 Outcomes considered by the NTWC to be critical or important for decision making: Neurocognitive disorders

Prioritised outcome domain	Measured with	Consensus rating	Results available for comparison 1 or 2?	Review ID							
				Lee 2011a	Robinson 2011	Strom 2016	Liu 2018	Hmwe 2019	O' Caoimh 2019	Margenfield 2019	Chen 2020a
Behavioural symptoms	NPI	Critical	Yes	--	--	--	?	--	--	✓	?
Functional capability	ADL	Critical	Yes	?	?	?	?	?	--	†	?
Cognitive function	MMSE	Critical	Yes	?	?	?	✓	?	--	✓	✓
Quality of life	Any validated measure	Critical	No	?	?	?	?	?	--	?	?
Psychosocial wellbeing	GDS (or other)	Important	No	?	?	?	?	--	--	?	?
Sleep quality	PSQI	Important	No	?	?	--	?	?	--	?	--

Abbreviations: ADL, activities of daily living; GDS, geriatric depression scale; MMSE, mini mental state examination; NPI, neuropsychiatric inventory, PSQI, Pittsburgh sleep quality index
✓ A study result is available for inclusion in the synthesis
† A study results is available for inclusion, but the systematic review only reports the direction of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.
-- The systematic review assessed this outcome but did not find or include any eligible primary studies that reported the outcome.
? The systematic review did not assess this outcome. It is unclear if the outcome was assessed by the primary studies included in the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

One systematic review (Margenfield 2019) identified two RCTs (Kwan 2017, Mariko 2015) comparing acupressure with sham in people living with dementia. One RCT contributed data relevant to 3 out of 6 critical or important outcomes. The other RCT did not measure, or report outcomes considered critical or important for this review.

Behavioural symptoms

There was one RCT (Mariko 2015) identified by Margenfield 2019 that reported behavioural symptoms (total 22 participants) measured with the neuropsychiatric inventory (NPI) at the end of treatment (4 weeks).

The NPI is used to assess neuropsychiatric symptoms in a variety of neurological conditions over the previous month (24). The questionnaire is completed by caregivers and includes 10 questions that examine 10 subdomains of behavioural functioning: delusions, hallucinations, agitation/aggression, dysphoria, anxiety, euphoria, apathy, disinhibition, irritability/lability, motor activity. There are 2 additional questions in a modified NPI that focus on night-time behavioural disturbances and appetite (24). Caregivers are asked to respond "yes" (present) or "no" (absent) to each question. If "yes", caregivers are asked to rate symptom frequency on a 4 point scale (1 = rarely, 2 = sometimes, 3 = often, 4 = very often) and symptom severity on a 3 point scale (1 = mind, 2 = moderate, 3 = severe). Caregiver distress associated with positive symptoms are rated on a 5 point scale (0 = no distress, 1 = minimal, 2 = mild, 3 = moderate, 4 = severe, 5 = extreme or very severe) (25).

The total NPI score can be calculated by adding the scores of the first 10 domain scores together. In most cases, the two neurovegetative items (night-time behavioural disturbances and appetite) are not included. If they are included, investigators must specify that the 12-item score is being used. The distress score is also not included in the total NPI score. For the 10-item scale, the composite NPI symptom score (frequency × severity) ranges from 0 (absence of behavioural symptoms) to 120 points (maximum severity). The total distress score is generated by adding together the scores of the first 10 (or all 12) items of the NPI distress questions, for a maximum score of 50 (or 60).

Higher scores indicate worse neuropsychiatric symptoms. For people with dementia, the MCID for the NPI is estimated to range between 2.77 and 3.18 points for symptom severity and 3.10 and 3.95 points for carer distress (26).

The review authors report mean change from baseline results from one RCT that suggest no difference between groups comparing acupressure with sham (MD –6.00; 95% –23.36, 11.36; p = 0.50). Similar results were observed comparing end of treatment results (MD –7.00; 95% –18.13, 4.13; p = 0.22).

Functional capability

There was one RCT (Mariko 2015) identified by Margenfield 2019 that reported functional capacity measured using the Barthel Index at the end of treatment (4 weeks). The Barthel Index is a widely used measure of functional disability and is used to measure the extent to which a person can function and mobilise independently during activities of daily living (27).

The review authors do not report any data for this outcome, but describe the results showed no significant difference between groups comparing acupressure with sham (p = not reported).

Cognitive function

There was one RCT (Mariko 2015) identified by Margenfield 2019 that reported cognitive function measured using the mini mental state examination (MMSE) at the end of treatment (4 weeks). The MMSE is a widely used test of cognitive function among the elderly, including testing of orientation, attention, memory, language and visual-spatial skills. Scores for each subdomain range from 0 (incorrect), 1 (correct), 6 (item administered, participant does not know answer), and 9 (test item not administered/unknown). For community dwelling older adults, the MCID is estimated to be a 5-point change (or less) over a five to ten year period (28).

The review authors report mean change from baseline results from one RCT that suggest no difference between groups comparing acupressure with sham (MD –0.00; 95% –7.24, 7.24; p = 1.00). End of treatment results also suggest there is no important difference between groups (MD –1.00; 95% –5.41, 3.41; p = 0.66).

Comparison 2 (vs control)

There were 4 RCTs (Feng 2015, Lin 2009, Sun 2016, Yang 2007) found by the included systematic reviews comparing acupressure with control (no intervention, waitlist, usual care) in people with neurocognitive disorders. One RCT (Sun 2016) assessing self-acupoint massage delivered as an adjunct to community services contributed data relevant to one out of 6 critical or important outcomes. The other RCTs did not measure, or report outcomes considered critical or important for this review.

One other RCT (Wan 2017) for which the comparator details were missing were also considered here.

Cognitive function

There were 2 RCTs (Sun 2016, Wan 2017) identified by the included systematic reviews that reported cognitive function measured using the mini mental state examination (MMSE) at the end of treatment (3/6 months[g] or not reported).

The MMSE is a widely used test of cognitive function among the elderly, including testing of orientation, attention, memory, language and visual-spatial skills. Scores for each subdomain range from 0 (incorrect), 1 (correct), 6 (item administered, participant does not know answer), and 9 (test item not administered/unknown). For community dwelling older adults, the MCID is estimated to be a 5-point change (or less) over a five to ten year period (28). The MCID at 3 to 6 months is unknown.

Available results from one RCT (Sun 2016) reported by Liu 2018 suggest an effect favouring the acupoint massage group when compared with no intervention (MD –3.10; 95% CI –3.92, –2.28; $p < 0.00001$).

Pooled results from 4 RCTs were reported by one systematic review (Chen 2020a) that suggested acupressure was effective in improving cognitive functioning in older adults (SMD 1.23; 95% CI 0.88, 1.59; p = not reported, $I^2 = 52.27\%$). Total participants were not reported. Two of the studies (Sun 2016, Wan 2017) were in people with dementia or mild cognitive impairment but data from the individual studies were incomplete and it was not possible to remove the 2 other studies from the analysis. The other 2 studies were in people with hypertensive heart disease (Lei 2015[h]) and sleep problems (Zeng 2016).

Comparison 3 (vs active)

There were 2 RCTs (Kwan 2017, Lin 2009) found by the included systematic reviews comparing acupressure with an active intervention in people with neurocognitive disorders. The RCTs did not measure, or report outcomes considered critical or important for this review.

S1.5.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There were 2 RCTs (Kwan 2017, Mariko 2015) found by the included systematic reviews comparing acupressure with sham in people living with dementia. One RCT contributed data relevant to 3 critical or important outcomes. The other RCT did not measure, or report outcomes considered critical or important for this review.

[g] The study was translated from Chinese. It is not clear if the self-treatment was for 3 months (with follow up at 6 months) or if treatment was for 6 months (with testing mid-treatment). Data at 6 months were reported.

[h] considered in the systematic review for shiatsu.

Acupressure compared to sham for Neurocognitive decline

Patient or population: Neurocognitive decline
Setting: community, institutionalised or in hospital
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Functional capacity assessed with: Barthel Index (higher is best) Scale from: 0 to 100 Follow-up: 4 weeks	No difference between groups (data not reported)		-	22 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on functional capacity in people with neurocognitive decline
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on health-related quality of life in people with neurocognitive decline is unknown
Behavioural symptoms assessed with: NPI (higher is worse) Scale from: 0 to 120 Follow-up: 4 weeks	The mean NPI score was **13 points**	MD **7 points lower** (18.13 lower to 4.13 higher)	-	22 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on behavioural symptoms in people with neurocognitive decline
Neurocognitive function assessed with: MMSE (higher is best) Scale from: 0 to 30 Follow-up: 6 months	-	MD **1.00 points lower** (5.41 lower to 3.41 higher)	-	23 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on neurocognitive function in people with neurocognitive decline **
Emotional wellbeing – not reported	-	-	-	(0 studies)	-	The effect of acupressure on emotional wellbeing in people with neurocognitive decline is unknown
Sleep quality – not reported	-	-	-	(0 studies)	-	The effect of acupressure on sleep quality in people with neurocognitive decline is unknown

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

** MCID is estimated to be a 5-point change (or less) over a five to ten year period (28). The MCID for a 3 to 6 months is unknown.

CI: confidence interval; **MD:** mean difference; **MMSE:** mini mental state exam; **NPI:** neuropsychiatric inventory

Acupressure compared to sham for Neurocognitive decline

Patient or population: Neurocognitive decline
Setting: community, institutionalised or in hospital
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. Inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with dementia or mild cognitive impairment and is probably generalisable to people with Alzheimer's disease with few caveats. Certainty of evidence not downgraded.
d. Very serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and no important difference). Certainty of evidence downgraded 2 levels.
e. Publication bias suspected. Evidence is limited to a small number of small trials. Certainty of evidence downgraded.

Comparison 2 (vs control)

There were 4 RCTs (Feng 2015, Lin 2009, Sun 2016, Yang 2007) found by the included systematic reviews comparing acupressure with control (no intervention, waitlist, usual care) in people with neurocognitive disorders. One RCT (Sun 2016) assessing self-acupoint massage delivered as an adjunct to community services contributed data relevant to one critical or important outcome. The other RCTs did not measure, or report outcomes considered critical or important for this review.

Acupressure compared to control (no intervention, waitlist, usual care) for Neurocognitive decline

Patient or population: Neurocognitive decline
Setting: community, institutionalised or in hospital
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Functional capacity – not reported	-	-	-	(0 studies)	-	The effect of acupressure on functional capacity in people with neurocognitive decline is unknown
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on health-related quality of life in people with neurocognitive decline is unknown

Acupressure compared to control (no intervention, waitlist, usual care) for Neurocognitive decline

Patient or population: Neurocognitive decline
Setting: community, institutionalised or in hospital
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Behavioural symptoms – not reported	-	-	-	(0 studies)	-	The effect of acupressure on behavioural symptoms in people with neurocognitive decline is unknown
Neurocognitive function assessed with: MMSE (higher is best) Scale from: 0 to 30 Follow-up: 4 weeks	The mean MMSE score was **25.3 points**	MD **3.1 points higher** (2.28 higher to 3.92 higher)	-	76 (1 RCT) # data from 1 RCT not included here	⊕⊕◯◯ LOW [a,b,c,d,e]	Acupressure may result in a slight improvement in neurocognitive function in people with neurocognitive decline **
Emotional wellbeing – not reported	-	-	-	(0 studies)	-	The effect of acupressure on emotional wellbeing in people with neurocognitive decline is unknown
Sleep quality – not reported	-	-	-	(0 studies)	-	The effect of acupressure on sleep quality in people with neurocognitive decline is unknown

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

** MCID is estimated to be a 5-point change (or less) over a five to ten year period (28).
# Data from one RCT (Wan 2017) not able to be included in the evidence synthesis. (SMD 0.95; 95% CI 0.49, 1.42) [number participants not reported)

CI: confidence interval; **MD:** mean difference; **MMSE:** mini mental state exam;

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. Inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with dementia or mild cognitive impairment who practice self-acupoint massage delivered as an adjunct to community services. It is probably generalisable to people with Alzheimer's disease with few caveats. Certainty of evidence not downgraded.
d. Serious imprecision. Single study. Optimal information size is probably not reached. Certainty of evidence downgraded.
e. Publication bias suspected. Evidence is limited to a small number of small trials. Certainty of evidence downgraded.

S1.5.5 Forest plots

Outcome results for people with neurocognitive decline (where additional analyses were required and able to be carried out) are presented in Figure S3 (behavioural symptoms) and Figure S4 (neurocognitive function).

Figure S3 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): Neurocognitive decline – behavioural symptoms (NPI)

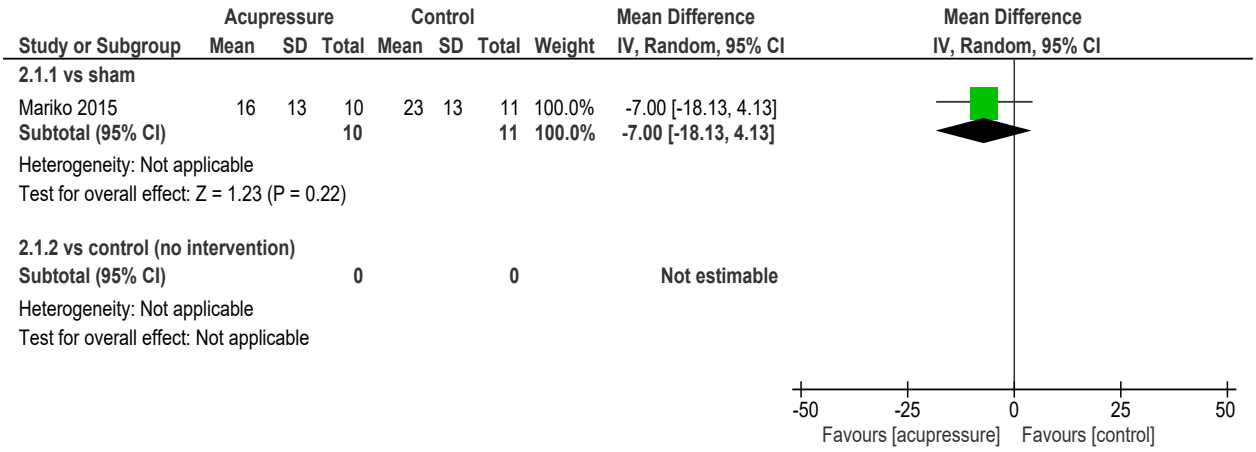
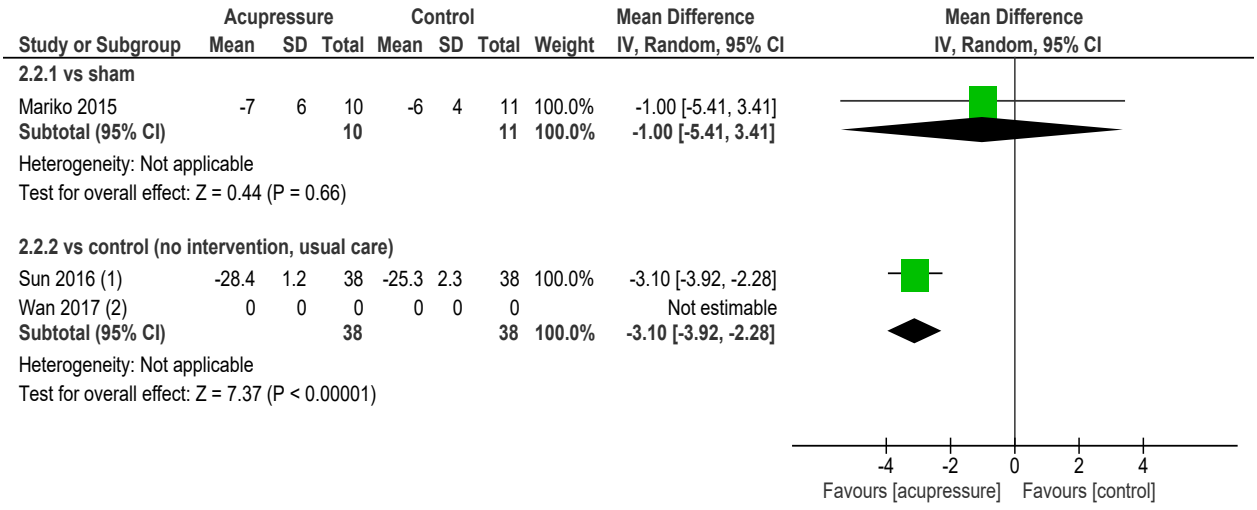


Figure S4 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): Neurocognitive decline – neurocognitive function (MMSE)



Footnotes

(1) Reported by Chen 2020a: SMD 1.69; 95% CI 1.17, 2.21

(2) Reported by Chen 2020a: SMD 0.95; 95% CI 0.49, 1.42 (Sample size not reported)

S1.6 Insomnia or sleep problems

S1.6.1 Description of studies

There were 15 citations (6, 11, 15, 20, 23, 29-38) corresponding to 15 systematic reviews (Cao 2009, Lee 2011c, Robinson 2011, Sarris 2011, Yeung 2012, Tan 2015, Hmwe 2016, Wang 2017, Capezuti 2018, Waits 2018, Harvie 2019, Hmwe 2019, Shang 2019, Chen 2020a, Samara 2020) identified in the literature that assessed acupressure compared to sham, control or an active intervention in people with insomnia or sleep disturbances. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are 2 systematic reviews awaiting classification (39, 40) (see Appendix C3.2) and 2 ongoing reviews (41) (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S6. Review details, including all outcome domains and measures, and the risk of bias (if available) of the included studies are provided in Appendix F1.2.

Most RCTs included by the systematic review authors were conducted in people with insomnia or sleep problems and are directly applicable to the population considered in the shiatsu evidence review. There were 9 RCTs (Chen 1999, Sun 2005, Hsu 2006, Nordio 2008, Reza 2010, Sun 2010, Abedian 2015, Lai 2017, Chen 2019) that compared acupressure with a sham intervention, with 5 RCTs (Chen 1999, Reza 2010, Lu 2013, Abedian 2015, Zeng 2016) also comparing to a control intervention (conversation, routine care, sleep hygiene advice). Two RCTs (Qui 1999, Zhou 2010) compared acupressure to an active intervention (benzodiazepines).

There were 6 other RCTs (He 2009, Lan 2009, Li 2007, Li 2009, Song 2007, Zhou 2007) found by one systematic review (Yeung 2012) that compared acupressure to an active intervention (typically benzodiazepines). The studies were conducted among inpatients or people attending outpatient clinics (no further details provided) and may not be applicable to the population included in the shiatsu review. No data were provided.

It was noted that there were several RCTs identified by the systematic review authors not included in their reviews due to inadequate reporting.

S1.6.2 Critical appraisal

Out of 14 systematic reviews, one review (Waits 2018) was judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest.

Eleven (11) reviews (Cao 2009, Lee 2011c, Robinson 2011, Yeung 2012, Tan 2015, Wang 2017, Capezuti 2018, Harvie 2019, Shang 2019, Chen 2020a, Samara 2020) had at least one critical flaw (i.e. did not meet, or partially meet, one of the prespecified critical AMSTAR-2 domains [4, 8, 9 or 11]). Of these, 5 systematic reviews (Samara 2020, Wang 2017, Tan 2015, Sarris 2011, Cao 2009) did not conduct a comprehensive literature search (domain 4), 2 (Tan 2015, Chen 2020a) failed to adequately describe the included studies in detail (domain 8), and one systematic review (Robinson 2011) did not use a satisfactory technique for assessing the risk of bias of individual studies (domain 9).

There were 8 systematic reviews (Yeung 2012, Wang 2017, Tan 2015, Shang 2019, Robinson 2011, Capezuti 2018, Harvie 2019, Lee 2011c) that did not perform a meta-analysis and one systematic review (Cao 2009) performed a meta-analysis but did not use appropriate methods to assess results (domain 11). One systematic review (Sarris 2011) did not meet any of the critical AMSTAR-2 domains (4, 8, 9 or 11).

A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

Table S6 List of included systematic reviews and overlap with eligible RCTs (per outcome): Insomnia

Review ID	Best available *	SR Outcome domains (measure)	Study ID												
			Chen 1999	Qui 1999	Sun 2005	Hsu 2006	Nordio 2008	Sun 2010	Reza 2010	Zhou 2010	Lu 2013	Abedian 2015	Zeng 2016	Lai 2017	Chen 2019
Cao 2009 (29)	†	Sleep quality (PSQI)	Y	--	--	--	--	--	--	--	--	--	--	--	--
		Increase in sleep time	!	Y	--	--	--	--	--	--	--	--	--	--	--
Lee 2011c (6)	†	Sleep quality (PSQI)	--	--	--	--	?	--	--	--	--	--	--	--	--
Robinson 2011 (15)	†	Sleep quality (PSQI)	?	--	--	?	--	?	?	--	--	--	--	--	--
Sarris 2011 (30)	X	Sleep quality (PSQI)	?	--	--	--	?	--	--	--	--	--	--	--	--
Yeung 2012 (31)^	†	Sleep quality (PSQI)	?	--	?	?	Y	!	Y	--	--	--	--	--	--
		Insomnia severity (AIS)	--	--	?	!	!	Y	!	--	--	--	--	--	--
Tan 2015 (32)	†	Sleep quality (PSQI)	--	--	--	?	?	!	?	--	--	--	--	--	--
		Insomnia severity (AIS)	--	--	--	--		?	--	--	--	--	--	--	--
Hmwe 2016	†	Sleep quality (PSQI) --	--	--	--	--	!	?	--	--	--	--	--	--	
		Insomnia severity (AIS) --	--	--	--	--	?	!	--	--	--	--	--	--	
Wang 2017 (34)	†	Sleep quality (PSQI)	--	--	--	--	Y	!	--	--	--	--	--	--	--
		Insomnia severity (AIS)	--	--	--	--	!	Y	--	--	--	--	--	--	--
Capezuti 2018 (35)	†	Sleep quality (PSQI)	?	--	--	--	--	!	?	--	--	--	--	--	--
		Insomnia severity (AIS)	!	--	--	--	--	?	--	--	--	--	--	--	--
Waits 2018 (36)	✓	Sleep quality (PSQI)	Y	--	--	--	Y	--	Y	Y	?	Y	--	--	--
Harvie 2019 (11)	†	Sleep quality (PSQI)	--	--	--	--	--	--	--	--	--	Y	--	--	--
Hmwe 2019 (20)	✓	Sleep quality (PSQI or AIS) ?	--	--	--	--	?	?	--	--	--	?	?	--	
		Neurocognitive function (MMSE) !	--	--	--	--	!	!	--	--	--	?	!	--	
		Psychosocial wellbeing !	--	--	--	--	!	!	--	--	--	!	!	--	
Shang 2019 (37)	†	Sleep quality (PSQI)	?	--	--	--	--	!	?	--	--	--	--	?	--
		Insomnia severity (AIS)	--	--	--	--	--	?	--	--	--	--	--	--	--
Chen 2020a (23)	†	Sleep quality (PSQI)	--	--	--	--	--	--	--	--	--	--	Y	--	--
		Neurocognitive function (MMSE)	--	--	--	--	--	--	--	--	--	--	Y	--	--
Samara 2020 (38)	†	Sleep quality (PSQI)	--	--	--	--	--	--	--	--	--	--	--	?	?

Abbreviations: AIS, Athens insomnia scale; PSQI, Pittsburgh sleep quality index

* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 8, domain 9 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])

✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

^ Six RCTs identified by Yeung 2012 not included in the overlap table. The studies reported 'effective rate' (not sleep quality).

Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]

? RCT is included in the systematic review, meets our PICO criteria but a study result is not available for the listed outcome measure [data is incomplete; result may be available in another SR]

! RCT is included in the systematic review but the SR indicates that the study does not measure the listed outcome [not measured]
-- RCT is not included in the systematic review

S1.6.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in people with insomnia are listed in Table S7.

Table S7 Outcomes considered by the NTWC to be critical or important for decision making: Insomnia

Prioritised outcome domain	Measured with	Consensus rating	Results available for comparison 1 or 2?	Review ID												
				Cao 2009	Lee 2011c	Robinson 2011	Sarris 2011	Yeung 2012	Tan 2015	Hmwe 2016	Wang 2017	Capezuti 2018	Waits 2018	Harvie 2019	Shang 2019	Chen 2020a	Samara 2020
Sleep quality	PSQI	Critical	Yes	*	*	X	*	*	*	*	*	*	✓†	*	X	†	†
Fatigue	NR	Critical	No	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Quality of life	NR	Critical	Yes	?	?	?	?	?	?	?	?	?	?	?	?	?	X
Cognitive function	MMSE	Critical	Yes	?	?	?	?	?	?	?	?	?	?	?	?	✓	?
Clinical effect	NR	Important	No	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Psychosocial wellbeing	GHQ-28	Important	Yes	?	?	?	?	?	?	?	✓	?	?	?	?	?	?
Cardiorespiratory	NR	Important	No	?	?	?	?	?	?	?	?	?	?	?	?	?	?

Abbreviations: NR, not reported; GHQ-28, 28-item general health questionnaire; PSQI, Pittsburgh sleep quality index
✓ A study result is available for inclusion in the synthesis
* A study result is available and reported in another systematic review nominated as the best available evidence.
X A study result is available for inclusion, but the systematic review reports incomplete data. Due to time and resource constraints, only the information presented in the systematic review is reported.
† A study result is available for inclusion, but the systematic review only reports the direction of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.
? The systematic review did not assess this outcome. It is unclear if the outcome was assessed by the primary studies included in the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

Five systematic reviews (Cao 2009, Yeung 2012, Wang 2017, Waits 2018, Samara 2020) included evidence from 8 RCTs comparing acupressure with a sham in people with insomnia or sleep problems and contribute data to 3 out of 7 critical or important outcomes.

Sleep quality

Three systematic review (Yeung 2012, Waits 2018, Samara 2020) included data from 8 RCTs (Chen 1999, Sun 2005, Hsu 2006, Nordio 2008, Reza 2010, Abedian 2015, Lai 2017, Chen 2019) reporting sleep quality (total 437 participants) measured with the Pittsburgh sleep quality index (PSQI) at the end of treatment (range 20 days to 8 weeks).

The PSQI is a 9-item questionnaire that assesses the sleep quality of an individual in the previous month. It assesses 7 sleep components including subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disorder (sleep fragmentation), use of sleeping medication and daytime dysfunction (42). Each item is scored from 0 to 3 with the total global score ranging from 0 (no problems) to 21 (severe problems). A score of 5 or more is associated with poor sleep quality (43). An MCID for the PSQI is not established for people with insomnia but reports range between 3.1 in adolescents (44) and 4.4 in patients after rotator cuff repair (45).

Pooled results from 4 RCTs (total 213 participants) (Chen 1999, Nordio 2008, Reza 2010, Abedian 2015) suggest an effect in favour of acupressure compared with sham groups (MD –3.20; 95% CI –4.10, –2.31; $p < 0.00001$; $I^2 = 26\%$). The clinical relevance of the observed difference is unclear, as participants in the intervention group continue to score above 5 points on the PSQI.

Data from another 2 RCTs (total 124 participants) (Lai 2017, Chen 2019) were not included in the analysis reported by Waits 2018, but combined data reported by another systematic review (Samara 2020) suggested an effect favouring acupressure (SMD –1.58; 95% CI –1.98, –1.17; $I^2 = 59\%$).

Data from 2 primary studies (Hsu 2006, Sun 2005) were not included in the meta-analysis by Waits 2018 (studies not identified). In another review (Yeung 2012) the authors did not include data from these 2 studies as they were judged to be at high risk of bias. However, authors described the results suggested an effect in favour of acupressure compared with sham groups (data not reported).

Quality of life

One systematic review (Samara 2020) included results from one RCT (Lai 2017) that reported quality of life measured SF-36 at the end of treatment (8 weeks).

The SF-36 is a multidimensional generic measure of health-related quality of life that comprises 36-items assessing eight domains: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning and mental health. Total scores for each domain are summarised on a scale from 0 (worse) to 100 (best) and are standardised to reflect a general population mean of 50 and standard deviation of 10 (46). The MCID for the SF-36 is estimated to be between 2 and 4 points (46).

The results from one study (total 62 participants) suggested an effect in favour of acupressure compared with the sham group (MD 5.09; 95% CI 1.38, 8.80; *p* = not reported).

Psychosocial wellbeing

One systematic review (Wang 2017) included results from one RCT (Nordio 2008) that reported psychosocial wellbeing measured using the 28-item general health questionnaire (GHQ-28) at the end of treatment (20 days).

The GHQ-28 is intended to screen for general (non-psychotic) mental health problems among primary care patients (47). Using a timeframe of "in the last two weeks", the tool consists of 28-items that measure concerns related to mental health across 4 domains (somatic symptoms, anxiety/insomnia, social dysfunction and severe depression). Responses are measured on a four-point scale (using a bimodal scoring method [0-0-1-1]), with higher scores indicating higher probability of psychiatric distress. Total scores that exceed 4 or 5 out of 28 suggest probable distress. The GHQ-28 is not designed to measure change over time therefore an MCID is not established.

The results from one study (total 40 participants) suggested an effect in favour of acupressure compared with the sham group (MD –1.41; 95% CI –2.77, –0.05; *p* = not reported). The clinical relevance of the observed change is unknown.

Comparison 2 (vs control)

Four systematic reviews (Cao 2009, Yeung 2012, Waits 2018, Chen 2020a) reported evidence from 6 RCTs comparing acupressure with control (no intervention, waitlist, or usual care) that were eligible for this comparison in people with insomnia (or sleep problems) and contributed data to 2 of 7 critical or important outcomes.

Sleep quality

There were 5 RCTs (Chen 1999, Reza 2010, Lu 2013, Abedian 2015, Zeng 2016) that measured sleep quality using the PSQI at the end of treatment (range 3 to 4 weeks or unknown). All RCTs suggested an effect in favour of acupressure compared with the control group, but data were incomplete, therefore were not able to be combined. The clinical importance of the observed difference is unclear.

One systematic review (Cao 2009) reported data from one RCT (Chen 1999) that suggested an effect in favour of acupressure compared with the control group (MD –6.32; 95% CI –7.47, –5.17; *p* = not reported).

One systematic review (Yeung 2012) reported data from one RCT (Reza 2010) which showed an effect in favour of acupressure compared with the control groups (MD –4.9; 95% CI –6.4, –3.3; $p < 0.001$).

One systematic review (Waits 2018) reported combined data from two RCTs (Chen 1999, Abedian 2015) (total 125 participants) that suggested an effect in favour of acupressure compared with the control group (conversation) (MD –5.13; 95% CI –5.86, –4.41; $p < 0.00001$; $I^2 = 0\%$). The review authors also reported data from 6 RCTs (total 363 participants) comparing acupressure with control (routine care) that also suggested an effect in favour of acupressure (MD –4.57; 95% CI –6.53, –2.60; $p < 0.00001$; $I^2 = 91\%$), but heterogeneity is high, with 4 out of the 6 RCTs being in people with sleep problems related to other underlying conditions (e.g. end-stage renal disease, hypertensive heart disease). Individual study data were not provided, therefore results from the 2 RCTs (Reza 2010, Lu 2013) included here could not be discerned.

One systematic review (Chen 2020a) reported pooled results from 6 RCTs (total participants not reported) that suggested acupressure was effective in improving sleep quality (SMD 0.85; 95% CI 0.49, 1.22; p = not reported, $I^2 = 68.73\%$). Only one RCT (Zeng 2016) was in people with insomnia or sleep problems (SMD 1.55; 95% CI 1.05 2.04), with the other RCTs considered with other population groups (hypertension, neurocognitive decline).

Cognitive function

There was one RCT (Zeng 2016) identified by one systematic review (Chen 2020a) that reported cognitive function measured using the mini mental state examination (MMSE) in people with insomnia or sleep problems at the end of treatment (timing not reported).

The MMSE is a widely used test of cognitive function among the elderly, including testing of orientation, attention, memory, language and visual-spatial skills. Scores for each subdomain range from 0 (incorrect), 1 (correct), 6 (item administered, participant does not know answer), and 9 (test item not administered/unknown). For community dwelling older adults, the MCID is estimated to be a 5-point change (or less) over a five to ten year period (28).

The systematic review reported pooled results from 4 RCTs (total participants not reported) that suggested acupressure was effective in improving cognitive functioning in older adults (SMD 1.23; 95% CI 0.88, 1.59; p = not reported, $I^2 = 52.27\%$); however, the data were incomplete, and it was not possible to remove the 3 other studies from the analysis. Results from one RCT (Zeng 2016) suggest an effect favouring acupressure (SMD 1.41; 95% CI 0.92, 1.89). The other 3 studies were in people with hypertensive heart disease (Lei 2015)[i] or people with neurocognitive impairment (Sun 2015, Wan 2017).

Comparison 3 (vs active)

Two systematic reviews (Cao 2009, Waits 2018) reported evidence from 2 RCTs comparing acupressure with an active intervention (benzodiazepines). One RCT (Zhou 2010) reported data relevant to one of 7 critical or important outcomes.

S1.6.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There were 8 RCTs found by the included systematic reviews comparing acupressure with sham in people with insomnia or sleep problems, of which 6 RCTs contributed data relevant to 3 critical or important outcomes.

[i] considered in the systematic review for shiatsu

Acupressure compared to sham for insomnia or sleep problems

Patient or population: insomnia (or sleep problems)
Setting: community or institutionalised
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Sleep quality assessed with: PSQI (higher is worse) Scale from: 0 to 21 Follow-up: range 20 days to 8 weeks	The mean PSQI score was **8.86 to 9.54 points**	MD **3.2 points lower** (4.10 lower to 2.31 lower)	-	213 (4 RCTs) # data from 4 RCTs (total 224 participants) not included here	⊕⊕⊕◯ MODERATE a,b,c,d,e	Acupressure probably results in a slight improvement on sleep quality in people with insomnia or sleep problems **
Fatigue – not reported	-	-	-	(0 studies)	-	The effect of acupressure on fatigue in people with insomnia or sleep problems is unknown
Quality of life assessed with: SF-36 (higher is better) Scale from: 0 to 100 Follow-up: 8 weeks	The mean score was **not reported**	MD **5.09 points higher** (1.38 higher to 8.80 higher)	-	62 (1 RCT)	⊕⊕◯◯ LOW a,b,c,e,f	Acupressure may result in an improvement in quality of life in people with insomnia or sleep problems ***
Neurocognitive function – not reported ##	-	-	-	(0 studies)	-	The effect of acupressure on neurcognitive function in people with insomnia or sleep problems is unknown
Global clinical improvement – not reported	-	-	-	(0 studies)	-	The effect of acupressure on global clinical improvement in people with insomnia or sleep problems is unknown
Psychosocial wellbeing assessed with: GHQ-28 (higher is worse) Scale from: 0 to 28 Follow-up: 20 days	The mean score was **not reported**	MD **1.41 points lower** (2.77 lower to 0.05 lower)	-	40 (1 RCT)	⊕⊕◯◯ LOW a,b,c,e,f	Acupressure may result in little to no improvement on psychosocial wellbeing in people with insomnia or sleep problems ****
Cardiorespiratory health – not reported	-	-	-	(0 studies)	-	The effect of acupressure on cardiorespiratory health in people with insomnia or sleep problems is unknown

Acupressure compared to sham for insomnia or sleep problems

Patient or population: insomnia (or sleep problems)
Setting: community or institutionalised
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				

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

** MCID not established in people with insomnia^. Ranges between 3.1 in adolescents (44) and 4.4 in patients after rotator cuff repair (45). A score of 5 or more on the PSQI is used as the threshold to diagnose sleep disturbance (i.e., is considered clinically relevant (43).
*** MCID is estimated to be between 2 and 4 points (46).
**** MCID is not established^.

^ In the absence of an MCID, the effect estimate was considered on 3 levels: small (MD <10% of the scale), moderate (MD between 10% to 20% of the scale), or large (MD more than 20% of the scale).

Data not able to be included for 2 studies (SMD –1.58; 95% CI –1.98, –1.17) and not available for 2 other studies.
Data from 1 RCT (Zheng 2016) included under neurocognitive decline

CI: confidence interval; **GHQ-28:** 28-item general health questionnaire; **MD:** mean difference; **PSQI:** Pittsburgh sleep quality index; **SF-36:** 36-item short form

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency or inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with insomnia or sleep problems, typically using self-acupressure. Certainty of evidence not downgraded.
d. No serious imprecision. Certainty of evidence not downgraded.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.
f. Serious imprecision. Wide confidence intervals (upper bound overlaps with no important difference). Certainty of evidence downgraded.

Comparison 2 (vs control)

There were 6 RCTs found by the included systematic reviews comparing acupressure with control (no intervention, waitlist or usual care) in people with insomnia or sleep problems that contributed data relevant to one critical or important outcomes.

Acupressure compared to control (no intervention, waitlist or usual care) for insomnia (or sleep problems)

Patient or population: insomnia (or sleep problems)
Setting: community or institutionalised
Intervention: acupressure
Comparison: control (no intervention, waitlist or usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Sleep quality assessed with: PSQI (higher is worse) Scale from: 0 to 21 Follow-up: range 20 days to 8 weeks	The mean PSQI score was **not reported**	MD **5.13 points lower** (range 5.86 lower to 4.41 lower)	-	125 (2 RCTs) # Data from 3 RCTs (total 111+ participants) not included here	⨁⨁⨁◯ MODERATE a,b,c,d,e	Acupressure probably results in a slight improvement on sleep quality in people with insomnia (or sleep problems) **
Fatigue – not reported	-	-	-	(0 studies)	-	The effect of acupressure on fatigue in people with insomnia or sleep problems is unknown
Quality of life– not reported	-	-	-	(0 studies)	-	The effect of acupressure on quality of life in people with insomnia or sleep problems is unknown
Cognitive function – not reported	-	-	-	(0 studies)	-	The effect of acupressure on cognitive function in people with insomnia or sleep problems is unknown
Global clinical improvement – not reported	-	-	-	(0 studies)	-	The effect of acupressure on global clinical improvement in people with insomnia or sleep problems is unknown
Psychosocial wellbeing – not reported	-	-	-	(0 studies)	-	The effect of acupressure on psychosocial wellbeing in people with insomnia or sleep problems is unknown
Cardiorespiratory health – not reported	-	-	-	(0 studies)	-	The effect of acupressure on cardiorespiratory health in people with insomnia or sleep problems is unknown

Acupressure compared to control (no intervention, waitlist or usual care) for insomnia (or sleep problems)

Patient or population: insomnia (or sleep problems)
Setting: community or institutionalised
Intervention: acupressure
Comparison: control (no intervention, waitlist or usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				

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

** MCID not established in people with insomnia. Ranges between 3.1 in adolescents (44) and 4.4 in patients after rotator cuff repair (45). A score of 5 or more on the PSQI is considered clinically relevant (43).

Data not able to be included for 3 studies that show an effect in favour of acupressure. One RCT (MD –4.9; 95% CI –6.4 to –3.3) and one RCT reported to show an effect in favour of acupressure, but data not reported.

CI: confidence interval; **GHQ-28:** 28-item general health questionnaire; **MD:** mean difference; **PSQI:** Pittsburgh sleep quality index; **SF-36:** 36-item short form

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

Explanations

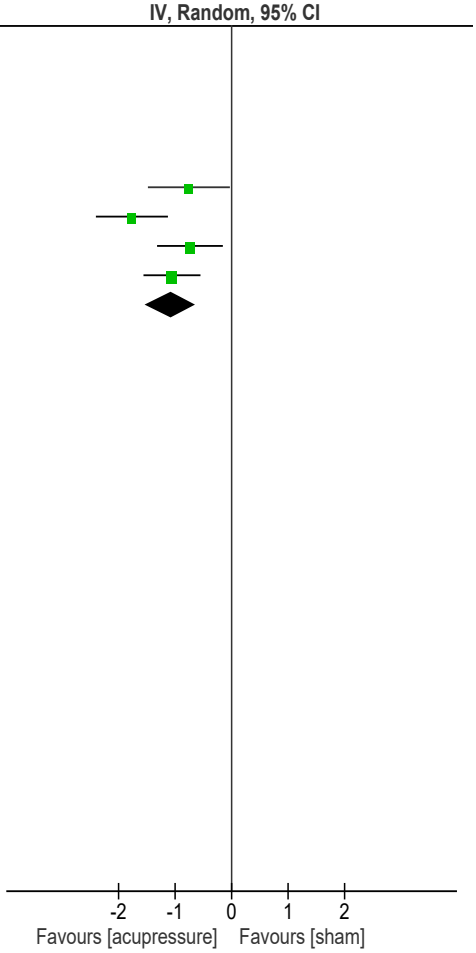
a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency. Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with insomnia or sleep problems, typically using self-acupressure. Certainty of evidence not downgraded.
d. No serious imprecision. Certainty of evidence not downgraded.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.

S1.6.5 Forest plots

Outcome results for people with insomnia or sleep disturbances (where additional analyses were required and able to be carried out) are presented in Figure S5 (sleep quality).

Figure S5 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): Insomnia – sleep quality

Study or Subgroup	Acupressure Mean	SD	Total	Sham Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
3.1.1 vs sham								
Sun 2005 (1)	0	0	25	0	0	25		Not estimable
Hsu 2006 (2)	0	0	25	0	0	25		Not estimable
Lai 2017 (3)	0	0	31	0	0	31		Not estimable
Chen 2019 (4)	0	0	31	0	0	31		Not estimable
Nordio 2008 (5)	6.61	2.97	18	8.86	2.82	15	21.0%	-0.76 [-1.47, -0.04]
Chen 1999 (6)	-5.93	2.36	28	-1.68	2.39	28	23.9%	-1.76 [-2.39, -1.14]
Reza 2010 (7)	6.84	2.79	25	9.54	4.25	26	26.0%	-0.74 [-1.31, -0.17]
Abedian 2015 (8)	-5.08	1.29	37	-2.19	3.63	36	29.1%	-1.06 [-1.55, -0.56]
Subtotal (95% CI)			**108**			**105**	**100.0%**	**-1.08 [-1.53, -0.63]**
Heterogeneity: Tau² = 0.12; Chi² = 6.83, df = 3 (P = 0.08); I² = 56%								
Test for overall effect: Z = 4.74 (P < 0.00001)								
3.1.2 vs control (conversation, usual care, sleep hygiene)								
Zeng 2016 (9)	0	0	0	0	0	0		Not estimable
Chen 1999 (10)	0	0	28	0	0	28		Not estimable
Lu 2013 (11)	0	0	30	0	0	30		Not estimable
Abedian 2015 (12)	0	0	37	0	0	32		Not estimable
Reza 2010 (13)	0	0	25	0	0	26		Not estimable
Subtotal (95% CI)			**120**			**116**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
3.1.3 vs active intervention (benzodiazepines)								
Zhou 2010 (14)	0	0	30	0	0	30		Not estimable
Subtotal (95% CI)			**30**			**30**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								



Footnotes

(1) Study identified by one systematic review. Authors report a significant effect favouring acupressure, but no data reported.
(2) Study was identified by 3 systematic reviews. Authors report a significant effect favouring acupressure, but no data reported
(3) Study identified by 2 systematic reviews. Authors report a significant effect favouring acupressure, but data not able to be included here.
(4) Study identified by one systematic review. Authors report a significant effect favouring acupressure, but data not able to be included here.
(5) Data reported by Waits 2018.
(6) Data reported by Waits 2018.
(7) Data reported by Waits 2018.
(8) Data reported by Waits 2018.
(9) Reported by Chen 2020a: SMD 1.41 (95% CI 0.92, 1.89) N = not reported
(10) Reported by Cao 2009: MD −6.32 (95% CI −7.47, −5.17)
(11) Study was identified by one systematic review. Authors report a significant effect favouring Acupressure, but no data reported
(12) Study was identified by one systematic review. Authors report a significant effect favouring Acupressure, but no data reported
(13) Reported by Yeung 2012: MD -4.9 (95% CI -6.4 to -3.3)
(14) Reported by Waits 2018: MD -2.40 (95% CI -4.48, -0.32) p = 0.02

S1.7 Hypertensive heart disease

S1.7.1 Description of studies

Five citations (20, 23, 33, 36, 48) corresponding to 5 systematic reviews (Hmwe 2016, Sibbritt 2018, Waits 2018, Hmwe 2019, Chen 2020a) were identified in the literature that assessed acupressure compared to sham, control or an active intervention in people with hypertensive heart disease. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are no systematic reviews awaiting classification (see Appendix C3.2) and no ongoing reviews (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S8. Review details, including all outcome domains and measures, and the risk of bias of the included studies are provided in Appendix F1.2.

The studies included by the systematic review authors had been conducted in people with hypertension and are directly applicable the population evaluated in shiatsu. One study (Lin 2016) compared acupressure with sham, 2 studies (Zheng 2014, Lei 2015) compared acupressure (or acupoint massage) with control (routine care, sleep hygiene education) and comparator details for two studies (Chen 2013, Li 2014) were not provided. No studies were found that compared acupressure to an active intervention in people with hypertension.

One study (Lei 2015) had been considered in the review for shiatsu (acupoint massage) therefore was not considered further.

Table S8 List of included systematic reviews and overlap with eligible RCTs (per outcome): Hypertensive heart disease

Review ID	Best available*	SR Outcome domains (measures)	Study ID				
			Chen 2013	Zheng 2014	Li 2014	~~Lei 2015~~	Lin 2016
Hmwe 2016 (33)	†	Sleepy quality (PSQI)	--	?	--	--	--
		Blood pressure (SBP, DBP)	--	?	--	--	--
Sibbritt 2018 (48)	†	Blood pressure (SBP, DBP)	--	--	--	--	?
Waits 2018 (36)	✓	Sleepy quality (PSQI)	--	?	--	?	--
Hmwe 2019 (20)	✓	Sleepy quality (PSQI)	--	--	--	?	--
Chen 2020a (23)	†	Sleepy quality (PSQI)	Y	--	Y	Y	--
		Cognitive function (MMSE)	!	--	!	Y	--

Abbreviations: DBP, diastolic blood pressure; ID, identification; PSQI, Pittsburgh sleep quality index; SBP, systolic blood pressure
~~Strikethrough~~ RCT was found but was considered in the evidence review for shiatsu (acupoint massage)
* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 8, domain 9 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])
✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]
? RCT is included in the systematic review, meets our PICO criteria but a study result is not available for the listed outcome measure [data is incomplete; result may be available in another SR]
! RCT is included in the systematic review but the SR indicates that the study does not measure the listed outcome [not measured]
-- RCT is not included in the systematic review

S1.7.2 Critical appraisal

Out of 5 systematic reviews, 2 reviews (Waits 2018, Hmwe 2019) were judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest. One review (Hmwe 2019) did not conduct a meta-analysis. The other 3 reviews (Chen 2020a, Hmwe 2016, Sibbritt 2018) had at least one critical flaw (i.e. did not meet, or partially met, one of the prespecified critical AMSTAR-2 domains [4, 8, 9 or 11]). Two systematic reviews (Hmwe 2016, Sibbritt 2018) did not conduct a comprehensive literature search (domain 4) and one systematic review (Chen 2020a) did not describe the included studies in adequate detail (domain 8).

A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.7.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in people with hypertensive heart disease are listed in Table S9.

Table S9 Outcomes considered by the NTWC to be critical or important for decision-making: Hypertensive heart disease

Prioritised outcome domain	Measured with	Consensus rating	Results available for comparison 1 or 2?	Review ID				
				Hmwe 2016	Waits 2018	Sibbritt 2018	Hmwe 2019	Chen 2020a
Cardiovascular health	Systolic and diastolic blood pressure	Critical	Yes	†	?	†	?	?
Cognitive function	MMSE	Important	No	?	?	?	?	--

Abbreviations: MMSE, mini mental state examination

† A study result is available for inclusion, but the systematic review only reports the direction of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.

-- The systematic review assessed this outcome but did not find or include any eligible primary studies that reported the outcome.

? The systematic review did not assess this outcome. It is unclear if the outcome was assessed by the primary studies. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

One systematic review (Sibbritt 2018) identified one RCT (Lin 2016) comparing acupressure with a sham in people with hypertensive heart disease that was eligible for this comparison and contributed data to one of the 2 critical or important outcomes.

Cardiovascular health

One systematic review (Sibbritt 2018) included one primary study (Lin 2016) (total 80 participants) that reported cardiovascular health measured by systolic and diastolic blood pressure (mmHg) at the end of treatment (15-30 mins after one treatment).

Systolic blood pressure (SBP) measures the force produced by the heart when it pumps blood out to the rest of the body. In the general adult population, an SBP below 120 mmHg is considered normal, whereas an SBP between 120 to 129 mmHg indicates high/elevated SBP (49). Diastolic blood pressure (DBP) measures the pressure in your arteries when the heart is at rest. In the general adult population, a DBP around 80 mmHg is considered normal, whereas a score between 85 to 89 mmHg indicates high/elevated DBP (49). The closer the score to 120/80 mmHg, the more stable the cardiorespiratory health.

The systematic review authors do not report any usable data but describe that in one study (Lin 2016) the results showed an effect in favour of acupressure compared with sham (p = not reported).

Comparison 2 (vs control)

There were 2 systematic reviews (Hmwe 2016, Waits 2018) that identified 1 RCT (Zheng 2014) comparing acupressure with control (routine care) in people with hypertensive heart disease that was eligible for this comparison and contributed data to 1 critical or important outcome.

Two other studies for which comparator details were not provided (Chen 2013, Li 2014) were to be considered here but did not report a critical or important outcome. The studies had been translated from Chinese by Chen 2020a and attempts to find additional information were unsuccessful.

Cardiovascular health

One systematic review (Hmwe 2016) included one primary study (Zheng 2014) (total 75 participants) that reported cardiovascular health measured by systolic and diastolic blood pressure (mmHg) at the end of treatment (4 weeks).

Systolic blood pressure (SBP) measures the force produced by the heart when it pumps blood out to the rest of the body. In the general adult population, an SBP below 120 mmHg is considered normal, whereas an SBP between 120 to 129 mmHg indicates high/elevated SBP (49). Diastolic blood pressure (DBP) measures the pressure in your arteries when the heart is at rest. In the general adult population, a DBP around 80 mmHg is considered normal, whereas a score between 85 to 89 mmHg indicates high/elevated DBP (49). The closer the score to 120/80 mmHg, the more stable the cardiorespiratory health.

The systematic review authors do not report any usable data but describe that in one study (Zheng 2014) the results showed an effect in favour of acupressure compared with control ($p < 0.001$).

Comparison 3 (vs active)

There were no systematic reviews found that identified any RCTs or NRSIs comparing acupressure with an active comparator in people with hypertensive heart disease.

S1.7.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There was one RCT (Lin 2016) identified by the included systematic reviews comparing acupressure with a sham in people with hypertensive heart disease that was eligible for this comparison, but reported data were incomplete.

Acupressure compared to sham for hypertensive heart disease

Patient or population: Hypertensive heart disease
Setting: Community or institutionalised
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Cardiovascular health assessed with: Systolic blood pressure (closer to 120 is best) Follow-up: immediate (15 to 30-minutes after)	Significant between-group effect reported but no data provided.		-	80 (1 RCT)	⊕◯◯◯ VERY LOW a,b,c,d,e	The evidence is very uncertain about the effect of acupressure on systolic blood pressure in people with hypertensive heart disease.

Acupressure compared to sham for hypertensive heart disease

Patient or population: Hypertensive heart disease
Setting: Community or institutionalised
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Cardiovascular health assessed with: Diastolic blood pressure (closer to 80 is best) Follow-up: immediate (15 to 30-minutes after)	Significant between-group effect reported but no data provided.		-	80 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on diastolic blood pressure in people with hypertensive heart disease.
Cognitive function – not reported	-	-	-	(0 studies)	-	The effect of acupressure on neurocognitive function in people with hypertensive heart disease is unknown

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

CI: confidence interval; **MD:** mean difference

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency or inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with hypertensive heart disease, typically using self-acupressure. Certainty of evidence not downgraded.
d. Very serious Imprecision. Imprecision not able to be assessed. Single study. Optimal information size is probably not reached. Certainty of evidence downgraded 2 levels.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.

Comparison 2 (vs control)

There was one RCT (Zheng 2014) identified by the included systematic reviews comparing acupressure with a control (routine care) in people with hypertensive heart disease that was eligible for this comparison, but reported data were incomplete.

Acupressure compared to control (no intervention, waitlist, usual care) for hypertensive heart disease

Patient or population: Hypertensive heart disease
Setting: Community or institutionalised
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Cardiovascular health assessed with: Systolic blood pressure (closer to 120 is best) Follow-up: 4 weeks	Significant effect reported but no data provided.		-	75 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on systolic blood pressure in people with hypertensive heart disease.
Cardiovascular health assessed with: Diastolic blood pressure (closer to 80 is best) Follow-up: 4 weeks	Significant effect reported but no data provided.		-	75 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on diastolic blood pressure in people with hypertensive heart disease.
Neurocognitive function – not reported	-	-	-	(0 studies)	-	The effect of acupressure on neurocognitive function in people with hypertensive heart disease is unknown

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

CI: confidence interval; **MD:** mean difference

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency or inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with hypertensive heart disease, typically using self-acupressure. Certainty of evidence not downgraded.
d. Very serious imprecision. Imprecision not able to be assessed. Single study. Optimal information size is probably not reached. Certainty of evidence downgraded 2 levels.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.

S1.8 Constipation

S1.8.1 Description of studies

Two citations (11, 23) corresponding to 2 systematic reviews (Harvie 2019, Chen 2020a) were identified in the literature that assessed acupressure compared to sham, control or an active intervention in people with constipation. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are no systematic reviews awaiting classification (see Appendix C3.2) and no ongoing reviews (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S10. Review details, including all outcome domains and measures, and the risk of bias of the included studies are provided in F1.2.

The studies included by the systematic review authors were conducted in adults with functional constipation and are directly applicable to the population evaluated in shiatsu. One study (Abbott 2014) compared acupressure (self-perineal) with control (no intervention). The comparator details for 3 studies (Wu 2012, Mo 2015, Liu 2017) were not provided. No studies were found that compared acupressure to an active intervention in people with functional constipation.

Table S10 List of included systematic reviews and overlap with eligible RCTs (per outcome): Constipation

Review ID	Best available*	SR Outcome domains (measures)	Study ID			
			Wu 2012	Abbott 2014	Mo 2015	Liu 2017
Harvie 2019 (11)	†	Quality of life (PAC-QoL, SF-12), BFI	--	?	--	--
Chen 2020a (23)	†	Bowel function (successful)	?	--	Y	?

Abbreviations: BFI, bowel function index; PAC-QoL, patient assessment of constipation - quality of life questionnaire; SF-12, 12-item short form

* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 8, domain 9 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])

✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]

? RCT is included in the systematic review, meets our PICO criteria but a study result is not available for the listed outcome measure [data is incomplete; result may be available in another SR]

-- RCT is not included in the systematic review

S1.8.2 Critical appraisal

No systematic reviews were judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest.

The 2 eligible reviews (Chen 2020a, Harvie 2019) had at least one critical flaw (i.e. did not meet, or partially met, one of the prespecified critical AMSTAR-2 domains [4, 8, 9 or 11]). A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.8.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in people with constipation are listed in Table S11.

Table S11 Outcomes considered by the NTWC to be critical or important for decision making: Constipation

Prioritised outcome domain	Measured with	consensus rating	Results available for comparison 1 or 2?	Review ID	
				Harvie 2019	Chen 2020a
Symptom severity	Bowel Function Index	Critical	Yes	X	?
Quality of life	PAC-QoL bowel function (disease specific preferred)	Critical	Yes	X	?
Clinical efficacy	Clinician-rated 'cure'	Critical	Yes	?	†

Abbreviations: PAC-QoL, patient assessment of constipation Quality of Life Questionnaire

X A study result is available for inclusion, but the systematic review reported incomplete data. Due to time and resource constraints, only the information presented in the systematic review is reported.

† A study result is available for inclusion, but the systematic review only reports the direction of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.

? The systematic review did not assess this outcome. It is unclear if the outcome was assessed by the primary studies included by the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

No systematic reviews were found that reported evidence from RCTs comparing acupressure with sham in people with functional constipation.

Comparison 2 (vs control)

Two systematic reviews (Chen 2020a, Harvie 2019) included evidence from 2 RCTs comparing acupressure with control in people with functional constipation that were eligible for this comparison and contributed data to 2 out of 3 critical or important outcomes. Data from 2 other RCTs were not available.

Symptom severity

One systematic review (Harvie 2019) identified one RCT (Abbott 2014) (total 100 participants) that reported symptom severity measured with the bowel function index (BFI) at the end of treatment (4 weeks).

The BFI is a self-reported questionnaire where participants rate 3-items (ease of defecation, feeling of incomplete bowel evacuation, and personal judgement of constipation) on a scale of 0 to 100 (higher means greater disease impact). A 12-point change in score is reported to be clinically important (50).

The review authors report that the RCT results[j] showed an effect in favour of acupressure compared with the control group, but data were incomplete (MD 13.8; 95% CI NR; $p < 0.01$).

Quality of life

One systematic review (Harvie 2019) identified one RCT (Abbott 2014) (total 100 participants) that reported quality of life measured with the Patient Assessment of Constipation – Quality of Life (PAC-QoL) questionnaire at the end of treatment (4 weeks).

The PAC-QoL is 28 item self-reported tool that is designed to measure the impact constipation has had on your daily life over the previous 2 weeks (51). There are 4 subscales (physical discomfort, psychosocial discomfort, worries and concerns, and satisfaction), each measured on a 5-point Likert scale, with the total PAC-QoL score reported on a scale of 0 (best) to 112 (worst). The first three subscales comprise the patient dissatisfaction index, with an overall score ranging from 0 to 96 (higher is worse). The satisfaction subscale includes four items that produce a combined score ranging from 0 to 16 (higher is better). A reduction of more than 1 in the total PAC-QoL score is suggested clinically meaningful in people with chronic non-cancer pain and opioid induced constipation (52).

The review authors report that the RCT results[jj] showed an effect in favour of acupressure compared with the control group but data were incomplete (MD 0.59; 95% CI NR; $p < 0.01$).

[j] Between group difference in mean change from baseline scores

Clinician-rated 'cure'

One systematic review (Chen 2020a) identified 3 RCTs (Wu 2012, Mo 2015, Liu 2017) that reported a 'return to successful bowel movement' at the end of treatment (timing not reported).

The review authors report results for one RCT (Mo 2015) (total participants not reported) that suggests an effect in favour of acupressure compared with the control group (SMD 0.46; 95% CI 0.00, 0.91; *p* = not reported). Results from 2 RCTs (Wu 2012, Liu 2017) were excluded from the meta-analysis due to substantial heterogeneity, but the systematic review authors do not describe the reason for exclusion and no other data were provided.

Comparison 3 (vs active)

There were no systematic reviews found that identified any RCTs or NRSIs comparing acupressure with an active comparator in people with functional constipation.

S1.8.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

No studies found.

Comparison 2 (vs control)

There were 2 RCTs found by the included systematic reviews comparing acupressure with control (no intervention, waitlist, usual care) in people with functional constipation that contributed data to 2 prioritised outcomes.

Acupressure compared to control (no intervention, waitlist, usual care) for functional constipation

Patient or population: functional constipation
Setting: community or hospital
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Symptom severity assessed with: BFI (higher is worse) Scale from: 0 to 100 follow-up: 4 weeks	The mean BFI score was not reported	MD **13.8 points lower** (CIs not reported)	-	100 (1 RCT)	⊕⊕◯◯ LOW [a,b,c,d,e]	Acupressure may result in reduced symptom severity in adults with functional constipation.**
Quality of life assessed with: PAC-QoL (higher is worse) Scale from: 0 to 112 follow-up: 4 weeks	The mean quality of life score was not reported	MD **0.59 points lower** (CIs not reported)	-	100 (1 RCT)	⊕⊕◯◯ LOW [a,b,c,d,e]	Acupressure may result in little to no improvement in quality of life in adults with functional constipation.***
Clinical efficacy assessed with: 'return to successful bowel movement' follow-up: not reported	-	SMD **0.46 SD lower^** (0.00 lower to 0.91 lower)	-	Not reported # (1 RCT) # missing data from 2 RCTs (total participant unknown)	⊕◯◯◯ VERY LOW [a,b,c,e,f]	The evidence is very uncertain about the effect of acupressure of clinical efficacy in adults with functional constipation.

Acupressure compared to control (no intervention, waitlist, usual care) for functional constipation

Patient or population: functional constipation
Setting: community or hospital
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				

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

^ As a rule of thumb, an SMD of 0.2 is considered a small difference, 0.5 is medium, and 0.8 is large difference (14).

** A 12-point change in score is reported to be clinically important (50).
*** A reduction of more than 1 in the total PAC-QoL score is suggested to be clinically meaningful in people with chronic non-cancer pain and opioid induced constipation (52).

BFI: bowel function index; **CI:** confidence interval; **MD:** mean difference; **PAC-QoL:** Patient Assessment of Constipation – Quality of life

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency or inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with functional constipation using perineal self-acupressure. It is not clear if this is applicable to the context of shiatsu. Certainty of evidence not downgraded.
d. Imprecision not able to be assessed. Single study. Optimal information size is probably not reached. Certainty of evidence downgraded.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.
f. Very serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and no importance difference). Certainty of evidence downgraded 2 levels.

S1.9 Chronic musculoskeletal pain

S1.9.1 Description of studies

Eight citations (6, 15, 53-58) corresponding to 8 systematic reviews (Lee 2011c, Robinson 2011, Kim 2012, Chen 2014, Yuan 2015, Yeganeh 2017, Godley 2020, Li 2021) were identified in the literature that assessed acupressure compared with sham, control or an active intervention in people with chronic musculoskeletal pain. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are no systematic reviews awaiting classification (see Appendix C3.2) and no ongoing reviews (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S12. Review details, including all outcome domains and measures, and the risk of bias of the included studies are provided in Appendix F1.2.

Table S12 List of included systematic reviews and overlap with eligible RCTs (per outcome): Chronic musculoskeletal pain

Review ID	Best available*	SR Outcome domains (measures)	Study ID													
			Salsali 2003	Lu 2004	Hsieh 2004	Hsieh 2006	Wang 2010	Zhang 2010	Zheng 2012	Wen 2015	Movahedi 2017	Zhang 2017	Liao 2018	Zhang 2018	~~Kobayashi 2019~~	Murphy 2019
Lee 2011c (6)	†	Pain (VAS, SF-MPQ)	--	--	?	?	--	--	--	--	--	--	--	--	--	--
Robinson 2011 (15)	X	Pain (VAS, SF-MPQ)	--	--	?	?	--	--	--	--	--	--	--	--	--	--
Kim 2012 (53)	✓	Pain (VAS, SF-MPQ)	--	--	Y	Y	--	--	--	--	--	--	--	--	--	--
		Disability (RMDQ, ODI)	--	--	!	Y	--	--	--	--	--	--	--	--	--	--
Chen 2014 (54)	†	Pain (VAS, SF-MPQ)	--	--	?	?	--	--	--	--	--	--	--	--	--	--
		Disability (RMDQ, ODI)	--	--	!	?	--	--	--	--	--	--	--	--	--	--
Yuan 2015 (55)	✓	Pain (VAS, SF-MPQ)	--	--	Y	Y	--	--	--	--	--	--	--	--	--	--
		Disability (RMDQ, ODI)	--	--	!	Y	--	--	--	--	--	--	--	--	--	--
Yeganeh 2017 (56)	✓	Pain (VAS)	?	--	--	--	--	--	--	--	--	--	--	--	--	--
Godley 2020 (57)	†	Pain (BPI, VAS)	--	--	--	?	--	--	--	--	?	--	--	--	--	?
		Disability (RMDQ)	--	--	--	?	--	--	--	--	!	--	--	--	--	?
		Fatigue (FSS, BFI)	--	--	--	!	--	--	--	--	?	--	--	--	--	?
Li 2021 (58) #	✓	Pain (VAS, SF-MPQ)	--	Y	Y	Y	!	Y	Y	Y	--	Y	!	!	Y	--
		Disability (RMDQ, ODI)	--	!	!	Y	!	!	!	!	--	!	!	!	Y	--
		Response rate	--	Y	!	!	Y	Y	Y	Y	--	Y	Y	Y	!	--

Abbreviations: BPI, brief pain inventory; BFI, brief fatigue inventory; FSS, fatigue severity score; ODI, Oswestry disability index; RMDQ, Roland-Morris disability score; SF-MPQ, short-form McGill pain questionnaire; VAS, visual analogue scale

~~Strikethrough~~ RCT was found but was considered in the evidence review for shiatsu.

* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 8, domain 9 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])

✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

Li 2021 included one other RCT (Chen 2015) that assessed acupressure in females with low back pain attributed to dysmenorrhoea. The study is considered in Section S1.10.

Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]

? RCT is included in the systematic review, meets our PICO criteria but a study result is not available [data is incomplete; result may be available in another SR]

! RCT is included in the systematic review, but the SR indicates that the study does not measure the listed outcome [not measured]
-- RCT is not included in the systematic review

The studies included by the systematic review authors were conducted in adults with chronic low back pain that was either non-specific or attributed to lumbar disc herniation and may not be directly applicable the populations evaluated in shiatsu (fibromyalgia, neck and shoulder stiffness). Two studies (Salsali 2003, Movahedi 2017) compared acupressure with a sham intervention and 3 studies (Salsali 2003, Kobayashi 2019, Murphy 2019) compared acupressure with control (no intervention or usual care).

There were 10 studies that compared acupressure with another intervention, being either physical therapy (Hsieh 2004, Hsieh 2006, Zhang 2017) or Tuina massage (Lu 2004, Wang 2010, Zhang 2010, Zheng 2012, Wen 2015, Liao 2018, Zhang 2018).

S1.9.2 Critical appraisal

Four systematic reviews (Kim 2012, Yuan 2015, Yegganeh 2017, Li 2021) that were judged to probably provide an accurate and comprehensive summary of the available studies to address the question of interest.

The other 4 eligible reviews (Lee 2011c, Robinson 2011, Chen 2014, Godley 2020) had at least one critical flaw (i.e. did not meet, or partially met, one of the prespecified critical AMSTAR-2 domains [4, 8, 9 or 11]). Of these, 2 systematic reviews (Chen 2014, Godley 2020) did not conduct a comprehensive literature search (domain 4), one review (Robinson 2011) failed to adequately describe the included studies in detail (domain 8), and 2 reviews (Robinson 2011, Godley 2020) did not use a satisfactory technique for assessing the risk of bias of individual studies (domain 9). There were 4 systematic reviews (Lee 2011c, Robinson 2011, Chen 2014, Godley 2020) that did not perform a meta-analysis. A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.9.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in people with chronic musculoskeletal pain are listed in Table S13.

Table S13 Outcomes considered by the NTWC to be critical or important for decision making: Chronic musculoskeletal pain

Prioritised outcome domain	Measured with	Consensus rating	Results available for comparison 1 or 2?	Review ID							
				Lee 2011c	Robinson 2011	Kim 2012	Chen 2014	Yuan 2015	Yegganeh 2017	Godley 2020	Li 2021
Pain	VAS or MPQ	Critical	Yes	*	*	†	*	*	†	†	✓
Functional capacity	SF-36 physical or ADL	Critical	No	?	?	?	?	--	?	?	?
Disability	ODI or other	Critical	Yes	?	?	?	*	*	--	†	✓
Quality of life	EQ-5D	Critical	No	?	?	?	?	--	?	?	?
Stress	PSS or other validated measure	Critical	No	?	?	?	?	?	?	?	?
Fatigue	Fatigue severity scale	Critical	Yes	?	?	?	?	?	?	†	?
Psychosocial wellbeing	STAI or SF-36 mental components	Critical	No	?	?	?	?	?	?	?	?

Abbreviations: ADL, activities of daily living; EQ-5D, Euro-quality of life 5-dimentions; MPQ, McGill pain questionnaire; ODI, Oswestry disability index; PSQI, Pittsburgh sleep quality index; SF-36, 36-item short form; STAI, state-trait anxiety index

✓ A study result is available for inclusion in the synthesis

* A study result is available and reported in another systematic review nominated as the best available evidence.

† A study result is available for inclusion, but the systematic review only reports the direction of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.

-- The systematic review assessed this outcome but did not find or include any eligible primary studies that reported the outcome.

? The systematic review did not assess this outcome. It is unclear is the outcome was assessed by the primary studies included in the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

Two systematic reviews (Godley 2020, Yeganeh 2017) identified 2 RCTs comparing acupressure with sham in people with chronic musculoskeletal pain that could have contributed data to 2 out of the 7 critical or important outcomes but the data were incomplete.

Pain

Two systematic reviews (Yeganeh 2017, Godley 2020) identified 2 RCTs (Salsali 2003, Movahedi 2017) (total 110 participants) that reported pain measured with a visual analogue scale (VAS) at the end of treatment (3 weeks).

The VAS is a subjective assessment of pain, reported by participants and measured on a continuous scale (mm) from 0 (no pain) to 100 (worst imaginable pain). Higher values indicate worse pain. An MCID for the pain VAS is reported to be around 20 mm in people with diabetic neuropathy, postherpetic neuralgia, chronic low back pain, fibromyalgia and osteoarthritis (59), but can vary from 8 to 40 mm across different patient groups (60).

The systematic review authors do not report any usable data but describe statistically significant results for both Salsali 2003 ($p < 0.0001$) and Movahedi 2017 ($p < 0.05$) that suggest an effect in favour of acupressure compared with a sham intervention.

Fatigue

One systematic review (Godley 2020) identified one RCT (Movahedi 2017) (total 50 participants) that reported fatigue measured with the fatigue severity scale at end of treatment (3 weeks). The 9-item fatigue severity scale assesses the impact fatigue has on certain activities of daily living. Items are rated on a 7-point scale form 1 (strongly disagree) to 7 (strongly agree). Total score ranges from 9 (no fatigue) to 63 (severe fatigue).

The systematic review authors do not report any usable data but describe results that suggest an effect in favour of acupressure compared with sham ($p < 0.05$).

Comparison 2 (vs control)

Two systematic reviews (Yeganeh 2017, Godley 2020) identified 2 RCTs comparing acupressure with control (no intervention or usual care) in people with chronic musculoskeletal pain that contributed data relevant to 3 out of 7 critical or important outcomes.

Pain

One systematic review (Godley 2020) identified one RCT (Murphy 2019) (total 67 participants) that reported pain measured with the Brief pain inventory (BPI) at the end of treatment (6 weeks). The systematic review authors do not report any usable data but describe results that suggest an effect in favour of acupressure compared with control ($p < 0.05$).

The BPI assesses pain severity and its interference on various aspects of life (including general activity, mood, sleep, mobility, activities of daily living, role-social, enjoyment). Each item is rated on a scale from 0 to 10, with the total scores calculated as an average of each item (score range 0 to 10). Higher scores mean worse pain. The 11-item measure can be reported as 2 subscales: pain severity (4-items) and pain interference (7-items). The MCID for the BPI in people with chronic low back pain is not established. In people with fibromyalgia it is estimated to be 2.2 points (61).

One other systematic review (Yeganeh 2003) identified 1 RCT (Salsali 2003) (total 60 participants) that measured pain (measure not reported) at the end of treatment (20 days). The systematic review authors do not report any usable data but describe results that suggest an effect in favour of acupressure compared with control (acetaminophen) ($p < 0.0001$).

One other systematic review (Li 2021) reported a meta-analysis involving 2 RCTs (Chen 2015, Kobayashi 2019) that are not included here. One RCT (Kobayashi 2019) was identified and already included in the shiatsu review (see main report) and one RCT (Chen 2015) is in females with dysmenorrhoea and is considered elsewhere (see Section S1.10).

Disability

One systematic review (Godley 2020) identified 1 RCT (Murphy 2019) (total 67 participants) that measured disability using the Roland-Morris Disability Questionnaire (RMDQ) at the end of treatment (6 weeks).

The RMDQ is a measure of how back pain affects functional activities in people with mild to moderate acute or chronic low back pain. Answers are scored on a range from 0 (no disability) to 24 (severe disability). There are also 18-item or 21-item versions. In people with chronic low back pain the minimal important difference is reported to be 5 points (62, 63), with an RMDQ threshold value of 4 (out of 24) suggested to identify those who met their goals compared with those who did not (64).

The systematic review authors do not report any usable data but describe results that suggest an effect in favour of stimulation acupressure compared with control ($p < 0.05$). The effect for relaxing acupressure compared with control was not significant (not data).

One systematic review (Li 2021) reported a meta-analysis involving 2 RCTs (Chen 2015, Kobayashi 2019) that reported disability measured with the Oswestry disability index (ODI) at the end of treatment (4 weeks). The results are not included here - one RCT (Kobayashi 2019) was identified and already included in the shiatsu review (see main report) and one RCT (Chen 2015) is in females with dysmenorrhoea and is considered elsewhere (see Section S1.10).

Fatigue

One systematic review (Godley 2020) identified 1 RCT (Murphy 2019) that reported fatigue measured with the brief fatigue inventory (BFI) at the end of treatment (6 weeks).

The BFI is designed to assess the severity and impact of cancer-related fatigue and is summarised on a scale from 0 (no fatigue) to 10 (as bad as you can imagine). In people with cancer, cut points for fatigue level suggested are 1–3 (mild), 4–7 (moderate), and 8–10 (severe), which corelate with functional interference, symptoms, depression, and QoL (65).

The systematic review authors do not report any usable data but describe results that suggest an effect in favour of stimulating acupressure compared with control ($p < 0.05$). The effect for relaxing acupressure compared with control was not significant (no data).

Comparison 3 (vs active)

There were 10 RCTs identified by the included systematic reviews that compared acupressure with an active intervention in people with chronic musculoskeletal pain, of which 7 RCTs contributed data to one out of the 7 critical or important outcomes. The other 3 RCTs did not measure or report a critical or important outcome.

Pain

One systematic review (Li 2021) included 3 RCTs (Hsieh 2004, Hsieh 2006, Zhang 2017) comparing acupressure with physical therapy and 4 RCT (Lu 2004, Zhang 2010, Zheng 2012, Wen 2015) comparing acupressure with tuina massage that reported pain intensity measured with a VAS at the end of treatment (range 20 days to 5 weeks).

The VAS is a subjective assessment of pain, reported by participants and measured on a continuous scale (mm) from 0 (no pain) to 100 (worst imaginable pain). Higher values indicate worse pain. An MCID for the pain VAS is reported to be around 20 mm in people with diabetic neuropathy, postherpetic neuralgia, chronic low back pain, fibromyalgia and osteoarthritis (59), but can vary from 8 to 40 mm across different patient groups (60).

The pooled results of 3 studies (total 335 participants) comparing acupressure with physical therapy suggest a greater reduction in pain in the acupressure group (SMD –0.88; 95% CI –1.10, –0.65; $p < 0.0001$; $I^2 = 29\%$ [fixed effect]). The pooled results of 4 studies (total 668 participants) comparing acupressure with tuina demonstrated an effect in favour of acupressure [SMD –1.92; 95% CI –3.09, –0.76; $p = 0.001$; $I^2 = NR$ [random effect]). Individual study results were not reported.

S1.9.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There were 2 RCTs found in the included systematic reviews comparing acupressure with sham in people with chronic musculoskeletal pain that contributed data to the outcomes of pain and fatigue. No evidence was found for other critical or important outcomes.

Acupressure compared to sham for chronic musculoskeletal pain

Patient or population: chronic musculoskeletal pain
Setting: community
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Pain assessed with: VAS (higher is worse) Scale from: 0 to 10 follow-up: 3 weeks	Significant between-group effect reported but no data provided.		-	110 (2 RCT)	⊕◯◯◯ VERY LOW a,b,c,d,e	The evidence is very uncertain about the effect of acupressure on pain in people with chronic musculoskeletal pain.
Functional capacity – not reported	-	-	-	(0 studies)	-	The effect of acupressure on functional capacity in people with chronic musculoskeletal pain is unknown
Disability – not reported	-	-	-	(0 studies)	-	The effect of acupressure on disability in people with chronic musculoskeletal pain is unknown
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on quality of life in people with chronic musculoskeletal pain is unknown
Stress – not reported	-	-	-	(0 studies)	-	The effect of acupressure on stress in people with chronic musculoskeletal pain is unknown
Fatigue assessed with: FSI (higher is worse) Scale from: 0 to 10 follow-up: 3 weeks	Significant between-group effect reported but no data provided.		-	50 (1 RCT)	⊕◯◯◯ VERY LOW a,b,c,d,e	The evidence is very uncertain about the effect of acupressure on fatigue in people with chronic musculoskeletal pain.
Psychosocial wellbeing – not reported	-	-	-	(0 studies)	-	The effect of acupressure on psychosocial wellbeing in people with chronic musculoskeletal pain is unknown

Acupressure compared to sham for chronic musculoskeletal pain

Patient or population: chronic musculoskeletal pain
Setting: community
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				

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

CI: confidence interval; **FSI:** fatigue severity index; **MD:** mean difference; **VAS:** visual analogue scale

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency or inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with chronic low back pain. Certainty of evidence not downgraded.
d. Very serious imprecision. Imprecision not able to be assessed. Optimal information size is probably not reached. Certainty of evidence downgraded 2 levels.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.

Comparison 2 (vs control)

There were 2 RCTs found in the included systematic reviews comparing acupressure with control (no intervention or usual care) in people with chronic musculoskeletal pain that contributed data to the outcomes of pain, disability and fatigue. No evidence was found for other critical or important outcomes.

Acupressure compared to control (no intervention, usual care) for chronic musculoskeletal pain

Patient or population: chronic musculoskeletal pain
Setting: community
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Pain assessed with: VAS (higher is worse) Scale from: 0 to 10 follow-up: 4 weeks	Significant between-group effect reported but no data provided.		-	127 (2 RCT)	⨁◯◯◯ VERY LOW a,b,c,d,e	The evidence is very uncertain about the effect of acupressure on pain in people with chronic musculoskeletal pain.
Functional capacity – not reported	-	-	-	(0 studies)	-	The effect of acupressure on functional capacity in people with chronic musculoskeletal pain is unknown

Acupressure compared to control (no intervention, usual care) for chronic musculoskeletal pain

Patient or population: chronic musculoskeletal pain
Setting: community
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Disability assessed with: ODI (higher is worse) Scale from: 0 to 100 follow-up: 4 weeks	Significant between-group effect reported but no data provided.		-	67 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on disability in people with chronic musculoskeletal pain.
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on quality of life in people with chronic musculoskeletal pain is unknown
Stress – not reported	-	-	-	(0 studies)	-	The effect of acupressure on stress in people with chronic musculoskeletal pain is unknown
Fatigue assessed with: FSI (higher is worse) Scale from: 0 to 10 follow-up: 3 weeks	Significant between-group effect reported for acupressure (stimulating) but not acupressure (relaxing). No data provided.		-	67 (1 RCT)	⊕◯◯◯ VERY LOW [a,b,c,d,e]	The evidence is very uncertain about the effect of acupressure on fatigue in people with chronic musculoskeletal pain.
Psychosocial wellbeing – not reported	-	-	-	(0 studies)	-	The effect of acupressure on Psychosocial wellbeing in people with chronic musculoskeletal pain is unknown

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

CI: confidence interval; **FSI:** fatigue severity scale; **MD:** mean difference; **ODI:** Oswestry disability index; **VAS:** visual analogue scale

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

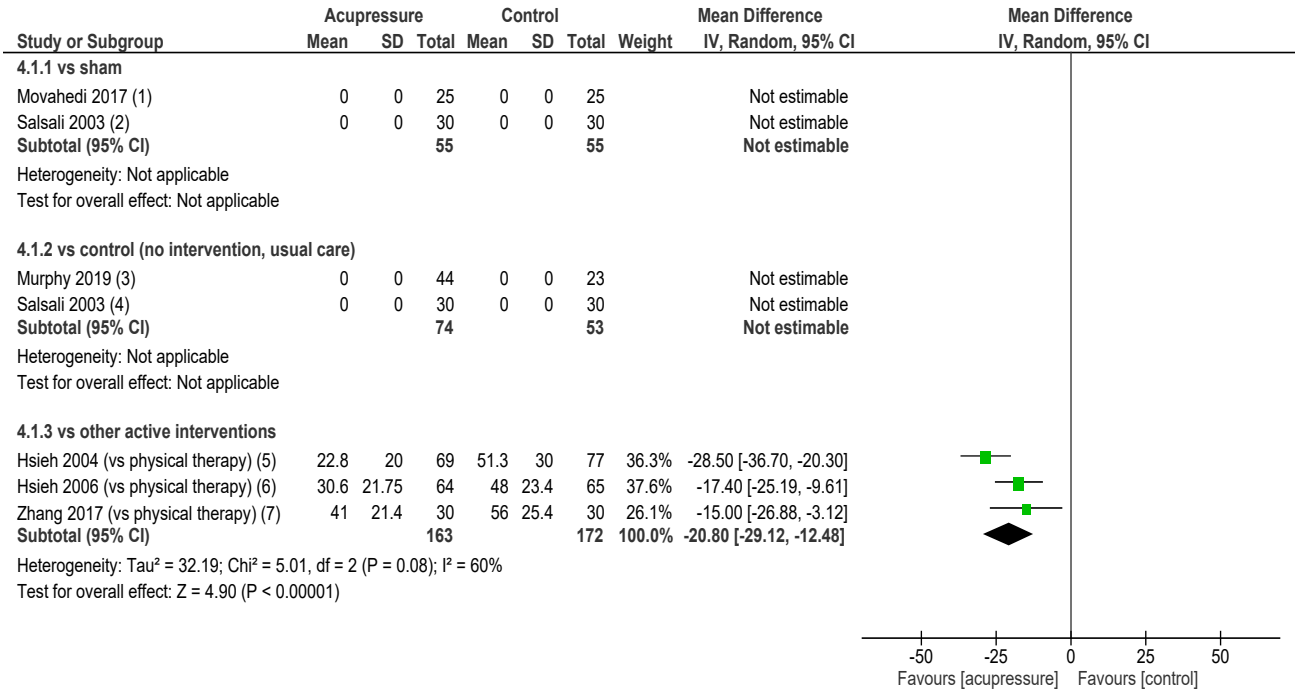
Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency or inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in people with chronic low back pain. Certainty of evidence not downgraded.
d. Very serious imprecision. Imprecision not able to be assessed. Optimal information size is probably not reached. Certainty of evidence downgraded 2 levels.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.

S1.9.5 Forest plots

Outcome results for people with chronic musculoskeletal pain (where additional analyses were required and able to be carried out) are presented in Figure S6 (pain).

Figure S6 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): chronic musculoskeletal pain – pain



Footnotes

(1) Data from Godley 2020: 62% reduction from baseline in the acupressure group; 44% reduction in the sham group ($p < 0.05$).
(2) Data from Yeganeh 2017: Authors report a significant effect favouring acupressure, but no data reported.
(3) Data from Godley 2020: 33% & 36% reduction from baseline in the acupressure groups (relaxing/stimulating); 9% reduction in the control group ($p < 0.05$).
(4) Data from Yeganeh 2017: Authors report a significant effect favouring acupressure, but no data reported.
(5) Data from Li 2021
(6) Data from Li 2021
(7) Data from Li 2021

S1.10 Dysmenorrhoea

S1.10.1 Description of studies

There were 18 citations (6, 11, 15, 32, 54, 58, 66-77) corresponding to 18 systematic reviews (White 2003, Cho 2010, Lathe 2011, Lee 2011c, Robinson 2011, Chung 2012, Chen 2013, Jiang 2013, Chen 2014, Kannan 2014, Abaraogu 2015, Song 2015, Tan 2015, Abaraogu 2016, Smith 2016, Armour 2019, Harvie 2019, Li 2021) identified in the literature that assessed acupressure compared to sham, control or an active intervention in females with primary dysmenorrhoea. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There is one systematic review awaiting classification (78) (see Appendix C3.2) and no ongoing reviews (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S14 (pain outcome) and Table S15 (outcomes other than pain). Review details, including all outcome domains and measures and the risk of bias of the included studies are provided in Appendix F1.2.

The RCTs included by the systematic review authors were conducted in females with primary dysmenorrhoea and are directly applicable the populations evaluated in shiatsu. Two studies (Charandabi 2011, Wong 2010) were cluster-randomised according to school dormitories. Six studies (Pouresmail 2002, Aghamiri 2005, Bazarganipour 2010, Kashefi 2010, Mirbagher-Ajorpaz 2011, Atrian 2013) compared acupressure with a sham intervention and 8 studies (Taylor 2002, Chen 2004, Chi 2004, Chen 2010, Wong 2010, Charandabi 2011, Chen 2015, Blodt 2018) compared acupressure with control (no intervention, rest or usual care). In 2 studies (Taylor 2002, Charandabi 2011) participants in both groups were permitted to use ibuprofen as needed.

There were 3 RCT that compared acupressure with another intervention, being either fish oil supplements (Zafari 2011), self-care exercises (Behbahani 2016) or ibuprofen (Poursemail 2002, Zafari 2011, Behbahani 2016).

S1.10.2 Critical appraisal

Of the 18 systematic reviews, 10 reviews (Cho 2010, Lathe 2011, Lee 2011c, Chen 2013, Jiang 2013, Kannan 2014, Abaraogu 2015, Smith 2016, Armour 2019, Li 2021) were judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest (i.e. met, or partially met, critical AMSTAR-2 domains (4, 8, 9 and 11).

The other 8 systematic reviews (White 2003, Robinson 2011, Chung 2012, Chen 2014, Song 2015, Tan 2015, Abaraogu 2016, Harvie 2019) had at least one critical flaw (i.e. did not meet, or partially meet one critical AMSTAR-2 domain [4, 8, 9 or 11]. Of these, 4 reviews (Chung 2012, Chen 2014, Tan 2015, Abaraogu 2016) did not justify publication restrictions (domain 4), 3 reviews (Robinson 2011, Song 2015, Tan 2015) failed to adequately describe the included studies in detail (domain 8), and 3 systematic reviews (White 2003, Robinson 2011, Harvie 2019) did not use a satisfactory technique for assessing the risk of bias of individual studies (domain 9). Several reviews did not perform a meta-analysis (White 2003, Cho 2010, Lathe 2011, Lee 2011c, Robinson 2011, Chen 2014, Song 2015, Tan 2015, Harvie 2019) (domain 11).

A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.10.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in primary dysmenorrhoea are listed in Table S16.

Table S14 List of included systematic reviews (reporting pain) and overlap with eligible RCTs: Primary dysmenorrhoea

Review ID	Best available*	SR Outcome domains (measures)	Study ID															
			Pouresmail 2002	Taylor 2002	Chen 2004	Chi 2004	Aghamiri 2005	Bazarganipour 2010	Chen 2010	Kashefi 2010	Wong 2010	Charandabi 2011	Mirbagher-Ajorpaz 2011	Zafari 2011	Atrian 2013	Chen 2015	Behbahani 2016	Blodt 2018
White 2003 (66)	†	Pain (VAS)	?	?	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Cho 2010 (67)	✓	Pain (VAS, SF-MPQ)	Y	Y	Y	--	Y	--		--	--	--	--	--	--	--	--	--
Lathe 2011 (68)	✓	Pain (VAS)	?	?	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Lee 2011c (6)	✓	Pain (VAS, MPQ)	?	?	?	--	--	--	--	--	--	--	--	--	--	--	--	--
Robinson 2011 (15)	†	Pain (VAS)	--	--	--	--	--	--	?	--	?	--	--	--	--	--	--	--
Chung 2012 (69)	†	Pain (VAS, SF-MPQ)	Y	--	Y	Y	--	Y	Y	Y	--	--	Y	--	--	--	--	--
Chen 2013 (70)	✓	Pain (VAS or SF-MPQ)	--	--	Y	--	--	--	--	Y	Y	--	Y	--	--	--	--	--
Jiang 2013 (71)	✓	Pain (VAS, AMS, SF-MPQ, MPQ)	Y	Y	Y	--	--	Y	Y	Y	Y	--	Y	--	--	--	--	--
Chen 2014 (54)	†	Pain (VAS, SF-MPQ)	--	--	--	--	--	--	?	?	?	--	?	--	--	--	--	--
Kannan 2014 (72)	✓	Pain (VAS)	Y	--	Y	--	--	--	Y	--	--	--	Y	--	--	--	--	--
Abaraogu 2015 (73)	✓	Pain (VAS, SF-MPQ)	Y	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Song 2015 (74)	†	Pain (VAS)	--	--	--	--	--	?	--	--	?	--	--	--	--	--	--	--
Tan 2015 (32)	†	Pain (VAS)	--	--	--	--	--	--	--	--	--	--	?	--	?	--	--	--
Abaraogu 2016 (75)	†	Pain (VAS, MPQ)	--	--	Y	--	--	--	--	Y	Y	--	Y	--	--	--	--	--
Smith 2016 (76)	✓	Pain (VAS)	a	--	Y	--	Y	Y	Y	Y	Y	!	Y	Y	--	--	--	--
Armour 2019 (77)	✓	Pain (VAS, AMS, SF-MPQ, MPQ)	Y	--	Y	--	Y	Y	Y	Y	Y	--	Y	Y	--	Y	Y	--
Harvie 2019 (11)	†	Pain (VAS)	--	--	--	--	--	--	--	--	--	--	--	--	--	--	?	?
Li 2021 (58)	✓	Pain (VAS)	--	--	--	--	--	--	--	--	--	--	--	--	--	Y	--	--

Abbreviations: AMS, Andersch and Milsom Scale; MPQ, McGill Pain Questionnaire; SF, short form; VAS, visual analogue scale
* Best available information (in any order) means the systematic review meets AMSTAR-2 domains 4, 8, 9 & 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])
✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11).
Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]
? RCT is included in the systematic review, meets our PICO criteria but a study result is not available [data is incomplete; result may be available in another SR]
! RCT is included in the systematic review, meets our PICO criteria but the SR indicates that the study does not measure the listed outcome [not measured]
-- RCT is not included in the systematic review
a. excluded by Smith 2016 as the details of randomisation could not be confirmed.

Table S15 List of included systematic reviews (outcomes other than pain) and overlap with eligible RCTs: Primary dysmenorrhoea

Review ID	Best available*	SR Outcome domains (measures)	Study ID															
			Pouresmail 2002	Taylor 2002	Chen 2004	Chi 2004	Aghamiri 2005	Bazarganipour 2010	Chen 2010	Kashefi 2010	Wong 2010	Charandabi 2011	Mirbagher-Ajorpaz 2011	Zafari 2011	Atrian 2013	Chen 2015	Behbahani 2016	Blodt 2018
Cho 2010 (67)	†	Symptom severity (SF-MDQ)	!	!	Y	--	!	--	--	--	--	--	--	--	--	--	--	--
Chung 2012 (69)	†	Symptom severity (SF-MDQ)	!	--	Y	!	--	!	Y	!	--	--	!	--	--	--	--	--
		Depression (BDI)	!	--	!	!	--	?	!	!	--	--	!	--	--	--	--	--
		Anxiety (VAS)	!	--	?	!	--	!	?	!	--	--	!	--	--	--	--	--
		BP, pulse, temperature	!	--	!	?	--	!	!	!	--	--	!	--	--	--	--	--
Jiang 2013 (71)	✓	Symptom severity (SF-MDQ, MDQ)	!	!	Y	--	--	!	Y	!	Y	--	!	--	--	--	--	--
		Anxiety (VAS)	!	!	?	--	--	!	?	!	!	--	!	--	--	--	--	--
Chen 2014 (54)	†	Symptom severity (SF-MDQ)	--	--	--	--	--	--	?	--	?	--	--	--	--	--	--	--
		Anxiety (VAS)	--	--	--	--	--	--	?	--	!	--	--	--	--	--	--	--
Song 2015 (74)	†	Symptom severity (MDQ, SF-MDQ)	--	--	--	--	--	!	--	--	?	--	--	--	--	--	--	--
Abaraogu 2016 (75)	†	Symptom severity (SF-MDQ)	--	--	Y	--	--	--	--	!	Y	--	!	--	--	--	--	--
		Anxiety (VAS)	--	--	Y	--	--	--	--	!	!	--	!	--	--	--	--	--
		Mental health (GHQ)	--	--	!	--	--	--	--	!	!	--	Y	--	--	--	--	--
		Quality of life	--	--	!	--	--	--	--	!	!	--	!	--	--	--	--	--
Smith 2016 (76)	✓	Symptom severity (MDQ, SF-MDQ)	--	--	Y	--	!	!	Y	!	Y	!	!	!	--	--	--	--
		medication use, ADL, QoL, absenteeism	a	--	!	--	!	!	!	!	!	!	!	!	--	--	--	--
Armour 2019 (77)	✓	Symptom severity (MDQ)	!	--	!	--	!	!	!	!	Y	Y	!	!	--	!	!	--

Abbreviations: ADL, activities of daily living; AMS, Andersch and Milsom Scale; BDI, becks depression inventory; GHQ, general health questionnaire; MDQ, Menstrual Distress Questionnaire; MPQ, McGill Pain Questionnaire; SF, short form; QoL, quality of life; VAS, visual analogue scale

* Best available information (in any order) means the systematic review meets AMSTAR-2 domains 4, 8, 9 & 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])

✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11).

Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]

? RCT is included in the systematic review, meets our PICO criteria but a study result is not available [data is incomplete; result may be available in another SR]

! RCT is included in the systematic review, meets our PICO criteria but the SR indicates that the study does not measure the listed outcome [not measured]

-- RCT is not included in the systematic review

a. excluded by Smith 2016 as the details of randomisation could not be confirmed.

Table S16 Outcomes considered by the NTWC to be critical or important for decision making: Primary dysmenorrhoea

Prioritised outcome domain	Measured with	Consensus rating	Results available for comparison 1 or 2?	Review ID																
				White 2003	Cho 2010	Lathe 2011	Lee 2011c	Robinson 2011	Chung 2012	Chen 2013	Jiang 2013	Chen 2014	Kannan 2014	Abaraogu 2015	Song 2015	Tan 2015	Abaraogu 2016	Smith 2016	Armour 2019	Harvie 2019
Symptom severity	Menstrual Symptom Severity List or similar	Critical	Yes	?	*	?	?	?	X	?	*	--	?	?	*	?	--	✓	#	?
Quality of life	SF-36 or similar	Critical	No	?	?	?	?	?	?	?	--	?	--	?	?	?	--	--	?	?
Pain	Visual Analogue Scale	Critical	Yes	*	*	*	*	†	*	*	†	*	*	*	*	X	*	✓	✓	†
Anxiety	Any validated measure	Critical	Yes	?	?	?	?	?	X	?	X	?	?	?	?	?	✓	?	?	?
Emotional function	General health questionnaire	Important	Yes	?	?	?	?	?	X	?	?	?	?	?	?	?	--	?	?	?
Depression	Any validated measure	Important	No	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Sleep quality	Any validated measure	Important	No	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

Abbreviations: SF-36, 36-item short form

✓ A study result is available for inclusion in the synthesis.

* A study result is available and reported in another systematic review nominated as the best available evidence.

X A study result is available for inclusion, but the systematic review reported incomplete data. Due to time and resource constraints, only the information presented in the systematic review is reported.

† A study result is available for inclusion, but the systematic review only reports the direction of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.

-- The systematic review assessed this outcome but did not find or include any eligible primary studies that reported the outcome.

? The systematic review did not assess this outcome. It is unclear if the outcome was assessed in the primary studies included in the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

The systematic review included relevant data within a meta-analysis that included primary studies not eligible for inclusion in this overview. Individual results from eligible primary studies were not available. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

The systematic reviews identified 6 RCTs that compared acupressure with a sham intervention (e.g. light touch, acupressure on non-acupoint) that contributed data to 2 out of 7 critical or important outcomes.

Pain

Several systematic reviews (Kannan 2014, Chung 2012, Song 2015, Tan 2015, Smith 2016, Armour 2019) identified 6 RCTs that reported pain intensity measured with a 10 cm visual analogue scale (VAS) (Pouresmail 2002, Aghamiri 2005, Kashefi 2010, Mirbagher-Ajorpaz 2011, Atrian 2013) or pain severity measured with the Andersch and Milsom scale (Bazarganipour 2010) at the end of treatment (after one to 3 menstrual cycles).

The VAS is a subjective assessment of pain, reported by participants and measured on a continuous scale (cm) from 0 (no pain) to 10 (worst imaginable pain). Higher values indicate worse pain. An MCID for the VAS is not available in females with primary dysmenorrhoea and can vary from 8 to 40 mm across different patient groups (mean 17 mm; 95% CI 15 to 19 mm) (60). It is estimated to be 10 mm (or 1 on a 10–point scale) in females with endometriosis (79).

Pooled results of 4 RCTs (total 380 participants) suggest an effect in favour of acupressure compared with the sham group (MD –2.23; 95% CI –3.61, –0.85; $p = 0.002$; $I^2 = 91\%$), but the heterogeneity is high. Data from one RCT (Atrian 2013) was not available to be used in the analysis, with the review authors (Tan 2015) noting that no difference was found between the acupressure and sham acupressure groups.

The Andersch and Milsom scale is used to assess the severity of menstrual pain and is measured on a four-point scale: no pain, mild pain without the need for painkillers, moderate pain that can be relieved by taking painkillers, and severe pain that does not go away with painkiller. Bazarganipour 2010 reported categorical data (i.e. number of participants who scored in each category) with pain severity reported to be lower in the acupressure group than in the control group at the fourth menstrual cycle ($p < 0.001$). One review (Chung 2012) reported these data as “cure rates”, recording the number pf participants who had a pain score of 0 at the end of treatment in each group (OR 0.211; 95% CI 0.107, 0.416; p = not reported).

Emotional functioning

One systematic review (Abaraogu 2016) included evidence from 1 RCT (Kashefi 2010) that reported mental wellbeing measured using the 28-item general health questionnaire (GHQ-28) at the end of treatment (2 menstrual cycles).

The GHQ-28 is intended to screen for general (non-psychotic) mental health problems among primary care patients (47). Using a timeframe of “in the last two weeks”, the tool consists of 28-items that measure concerns related to mental health across 4 domains (somatic symptoms, anxiety/insomnia, social dysfunction and severe depression). Responses are measured on a four-point scale (using a bimodal scoring method [0-0-1-1]), with higher scores indicating higher probability of psychiatric distress. Total scores that exceed 4 or 5 out of 28 suggest probable distress. The GHQ-28 is not designed to measure change over time therefore an MCID is not established.

The results from one study (total 86 participants) suggest an effect that favours acupressure when compared with the control group (MD –3.58; 95% CI –4.71, –2.45; $p < 0.00001$).

Comparison 2 (vs control)

The systematic reviews identified 8 RCTs comparing acupressure with control (no intervention, usual care) in people with primary dysmenorrhoea that contributed data to 3 of the 7 critical or important outcomes.

Symptom severity

Two systematic reviews (Smith 2016, Armour 2019) included evidence from 4 RCTs (Chen 2004, Chen 2010, Wong 2010, Charandabi 2011) that reported symptom severity measured with the menstrual distress questionnaire (MDQ) at the end of treatment (after one to 3 menstrual cycles).

The MDQ is a 46 and 47-item scale, reported by participants and measured on a continuous scale from 0 (not at all disabling) to 6 (partially disabling). There are 8 subgroups of symptoms relating to pain, concentration, behavioural changes (e.g. school avoidance), autonomic reactions (e.g., faint, cold sweats), water retention, negative affect (e.g., crying, loneliness, depression), arousal (e.g., bursts of energy, excitement), and control (e.g., heart pounding, feeling of suffocation, fuzzy vision). Higher values indicate worse symptom severity, with scores above 49 points suggesting very acute symptoms. Scores below 16 indicate minor symptoms, scores between 17 and 32 indicate moderate symptoms, and scores between 33 and 48 indicate acute severity (80). An MCID for the MDQ has not been established.

Pooled results from 3 RCTs (total 160 participants) suggest there is no difference between the acupressure group compared with the control group (MD –1.93; 95% CI –5.57, 1.70; $p = 0.30$; $I^2 = 81\%$). Heterogeneity was high and the direction of treatment effect was inconsistent.

Smith 2016 noted that data from one RCT (Charandabi 2011) were not able to be used due to data being presented within subcategories of dysmenorrhoea. The study was reported by Armour 2019 to measure menstrual symptoms severity on a 5-point Likert scale for 8 symptoms (cramp, headache, back pain, leg pain, depression, irritability, general pain and abdominal pain), noting that a greater reduction in symptoms was observed in the acupressure group (not data provided).

Pain

Several systematic reviews (Chung 2012, Jiang 2013, Abaraogu 2016, Smith 2016, Armour 2019, Harvie 2019, Li 2021) included evidence from 7 primary studies (Taylor 2002, Chen 2004, Chi 2004, Chen 2010, Chen 2015, Wong 2010, Blodt 2018) that reported pain measured with a VAS or numeric pain rating scale (NPRS) at the end of treatment (range end of 2 cycles to 12 months).

The VAS is a subjective assessment of pain, reported by participants and measured on a continuous scale (cm) from 0 (no pain) to 10 (worst imaginable pain). Higher values indicate worse pain. An MCID for the VAS is not available in people with primary dysmenorrhoea and can vary from 8 to 40 mm across different patient groups (mean 17 mm; 95% CI 15 to 19 mm) (60). It is estimated to be 10 mm (or 1 on a 10–point scale) in females with endometriosis (79).

Pooled results from 5 studies (total 383 participants) suggest an effect that favours acupressure when compared with the control group (MD –1.49; 95% CI –2.61, –0.37; $p = 0.009$; $I^2 = 90\%$). Statistical heterogeneity was high.

One systematic review (Chung 2012) included one additional RCT (Chi 2004) that measured pain immediately after treatment. The pain measure used was not reported by the systematic review authors. The systematic review authors do not report any usable data but describe that the results showed an effect in favour of acupressure compared with control.

One systematic review (Harvie 2019) included one additional RCT (Blodt 2018) that measured pain with a numerical rating scale (0 to 10) at the end of treatment (6 cycles). The review noted a greater reduction in pain in the acupressure group compared with control (MD –1.4; 95% CI –2.0, –0.8; $p < 0.001$) but data were incomplete and not able to be added to the meta-analysis.

Anxiety

One systematic review (Abaraogu 2016) included evidence from 1 RCT (Chen 2004) that reported anxiety measured with a visual analogue scale at the end of treatment (2 menstrual cycles). One other RCT (Chen 2010) is reported by various systematic reviews to measure anxiety, but no data were provided.

The VAS-A is a subjective assessment of anxiety, reported by participants and measured on a continuous scale (cm) from 0 (no anxiety) to 10 (worst imaginable anxiety). Higher values indicate worse anxiety. An MCID for the VAS-A is not available.

The results from one study (total 69 participants) suggest no difference between the treatment groups comparing acupressure with control (MD 0.50; 95% CI –0.54, 1.54; $p = 0.34$).

Comparison 3 (vs active)

The systematic reviews identified 3 RCTs comparing acupressure with another intervention in people with primary dysmenorrhoea that contributed data to 1 of the 7 critical or important outcomes.

Pain

Three systematic reviews (Jiang 2013, Smith 2016, Armour 2019) included evidence from 3 RCTs (Pouresmail 2002, Behbahani 2016, Zafari 2011) that reported pain intensity measured with a VAS or NPRS at the end of treatment (range one to 3 menstrual cycles).

Pooled results from 3 RCTs (total 360 participants) reported by Armour 2019 suggest an effect that favours the control group (ibuprofen) (Hedges *g*: 0.759; 95% CI 0.145, 1.373, *p* = 0.015).

Data for the other comparator groups (exercise, fish oil) included in 2 RCTs (Zafari 2011, Behbahani 2016) were not provided.

S1.10.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There were 6 RCTs found by the included systematic reviews that compared acupressure with a sham intervention and contributed data to 2 out of 7 critical or important outcomes. No evidence was found for other critical or important outcomes.

Acupressure compared to sham for dysmenorrhoea

Patient or population: dysmenorrhoea
Setting: community
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Symptom severity – not reported	-	-	-	(0 studies)	-	The effect of acupressure on symptom severity in people with dysmenorrhoea is unknown
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on quality of life in people with dysmenorrhoea is unknown
Pain assessed with: VAS (higher is worse) Scale from: 0 to 10 follow-up: range one to 3 menstrual cycles	The mean pain score was **5 points**	MD **2.23 points lower** (3.61 lower to 0.85 lower)	-	380 (4 RCTs) # missing data from 2 RCTs (238 participants)	⊕⊕◯◯ LOW [a,b,c,d]	Acupressure may reduce pain in people with dysmenorrhoea.**
Anxiety – not reported	-	-	-	(0 studies)	-	The effect of acupressure on anxiety in people with dysmenorrhoea is unknown
Emotional function assessed with: GHQ-28 (higher is worse) Scale from: 0 to 28 follow-up: 2 menstrual cycles	The mean GHQ-28 score was **7.57 points**	MD **3.58 points lower** (4.71 lower to 2.45 lower)	-	86 (1 RCT)	⊕◯◯◯ VERY LOW [a,c,e,f,g]	The evidence is very uncertain about the effect of acupressure on emotional functioning in people with dysmenorrhoea.***

Acupressure compared to sham for dysmenorrhoea

Patient or population: dysmenorrhoea
Setting: community
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Depression – not reported	-	-	-	(0 studies)	-	The effect of acupressure on depression in people with dysmenorrhoea is unknown
Sleep quality – not reported	-	-	-	(0 studies)	-	The effect of acupressure on sleep quality in people with dysmenorrhoea is unknown

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

** MCID is estimated to be 1.0 cm in females with endometriosis (79).
*** Total scores that exceed 4 or 5 out of 28 suggest probable distress (47).

1 RCT suggests an effect favouring acupressure, 1 RCT suggests no difference between groups.

CI: confidence interval; **GHQ-28:** 28-item general health questionnaire; **MD:** mean difference; **VAS:** visual analogue scale

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. Serious inconsistency. Point estimates vary widely and confidence intervals do not overlap for some studies. Substantial statistical heterogeneity ($I^2 > 90\%$) that cannot be explained. Certainty of evidence downgraded.
c. No serious indirectness. The available evidence in people with dysmenorrhoea. Certainty of evidence not downgraded.
d. Serious imprecision. Wide confidence intervals (lower bound overlaps with no important difference). Certainty of evidence downgraded.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.
f. Inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
g. Very serious imprecision. Single study with wide confidence intervals (lower bound overlaps with no important difference). Certainty of evidence downgraded 2 levels.

Comparison 2 (vs control)

There were 8 RCTs identified by the included systematic reviews comparing acupressure with control (no intervention, usual care) in people with primary dysmenorrhoea that contributed data to 3 of the 7 critical or important outcomes.

Acupressure compared to control (no intervention, waitlist, usual care) for dysmenorrhoea

Patient or population: dysmenorrhoea
Setting: community
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Symptom severity Assessed with: MDQ (higher is worse) Scale from: 0 to ? Follow-up: range one to 3 menstrual cycles	The mean MDQ score was **24.8 points**	MD **1.93 points lower** (5.57 lower to 1.70 higher)		380 (3 RCTs) # missing data from 1 RCT (72 participants)	⨁⨁◯◯ LOW [a,b,c,d]	Acupressure may result in little to no effect on symptom severity in people with dysmenorrhoea. **
Quality of life – not reported	-	-	-	(0 studies)	-	The effect of acupressure on quality of life in people with dysmenorrhoea is unknown
Pain intensity assessed with: VAS (higher is worse) Scale from: 0 to 10 follow-up: range one to 3 menstrual cycles	The mean pain score was **4.7 points**	MD **1.49 points lower** (2.61 lower to 0.37 lower)	-	363 (5 RCTs) ## missing data from 2 RCTs (281 participants)	⨁⨁◯◯ LOW [a,b,c,d]	Acupressure may result in a reduction in pain intensity in people with dysmenorrhoea. ***
Anxiety assessed with: VAS (higher is worse) Scale from: 0 to 10 follow-up: 2 menstrual cycles	The mean anxiety score was **2.76 points**	MD **0.50 points higher** (0.54 lower to 1.54 higher)	-	69 (1 RCT) missing data from 1 RCT (71 participants)	⨁◯◯◯ VERY LOW [a,c,e,f,g]	The evidence is very uncertain about the effect of acupressure on anxiety in people with dysmenorrhoea.
Emotional function – not reported	-	-	-	(0 studies)	-	The effect of acupressure on emotional function in people with dysmenorrhoea is unknown
Depression – not reported	-	-	-	(0 studies)	-	The effect of acupressure on depression in people with dysmenorrhoea is unknown

Acupressure compared to control (no intervention, waitlist, usual care) for dysmenorrhoea

Patient or population: dysmenorrhoea
Setting: community
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Sleep quality – not reported	-	-	-	(0 studies)	-	The effect of acupressure on sleep quality in people with dysmenorrhoea is unknown

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

** MCID not established. Scores between 17 and 32 indicate moderate symptoms (80)
*** MCID is estimated to be 1.0 cm in females with endometriosis (79).

1 RCT with incomplete data that suggest a greater reduction in symptom severity in the acupressure group compared with control.
2 RCTs with incomplete data that suggest a greater reduction in pain in the acupressure group compared with control.

CI: confidence interval; **MD:** mean difference; **MDQ:** menstrual distress questionnaire; **VAS:** visual analogue scale

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

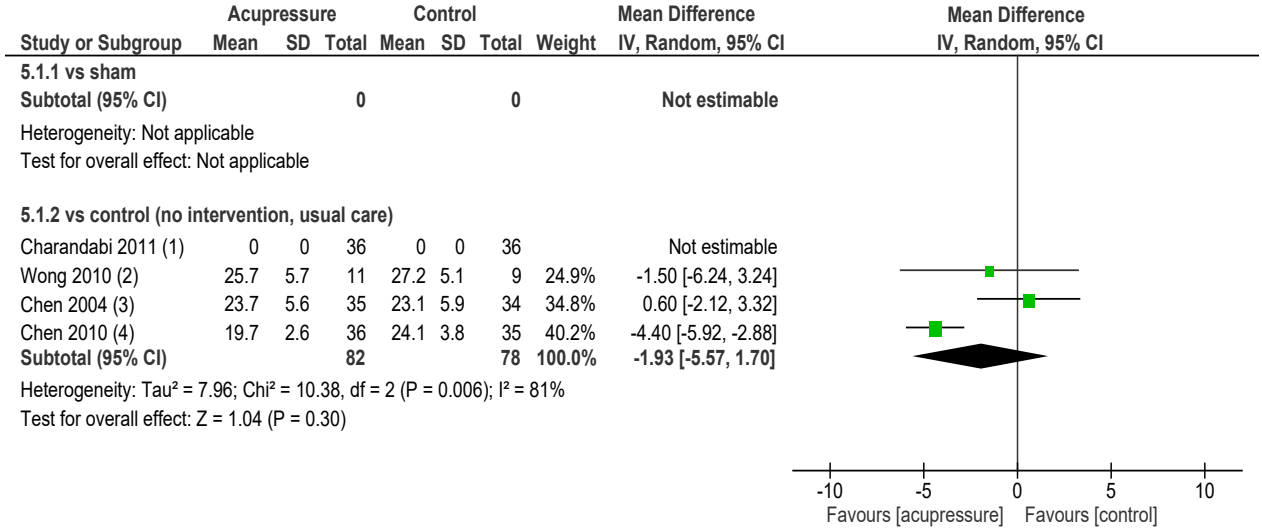
Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. Serious inconsistency. Point estimates vary widely and confidence intervals do not overlap for some studies. Substantial statistical heterogeneity (I^2 > 80%) that cannot be explained. Certainty of evidence downgraded.
c. No serious indirectness. The available evidence in people with dysmenorrhoea. Certainty of evidence not downgraded.
d. Serious imprecision. Wide confidence intervals (lower bound overlaps with no important difference). Certainty of evidence downgraded.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.
f. Inconsistency not able to be assessed (1 study). Certainty of evidence not downgraded.
g. Very serious imprecision. Single study with wide confidence intervals (lower bound overlaps with no important difference). Certainty of evidence downgraded 2 levels.

S1.10.5 Forest plots

Outcome results for people with dysmenorrhoea (where additional analyses were required and able to be carried out) are presented in Figure S7 (symptom severity), Figure S8 (pain), Figure S9 (anxiety) and Figure S10 (emotional functioning).

Figure S7 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): dysmenorrhoea – symptom severity



Footnotes

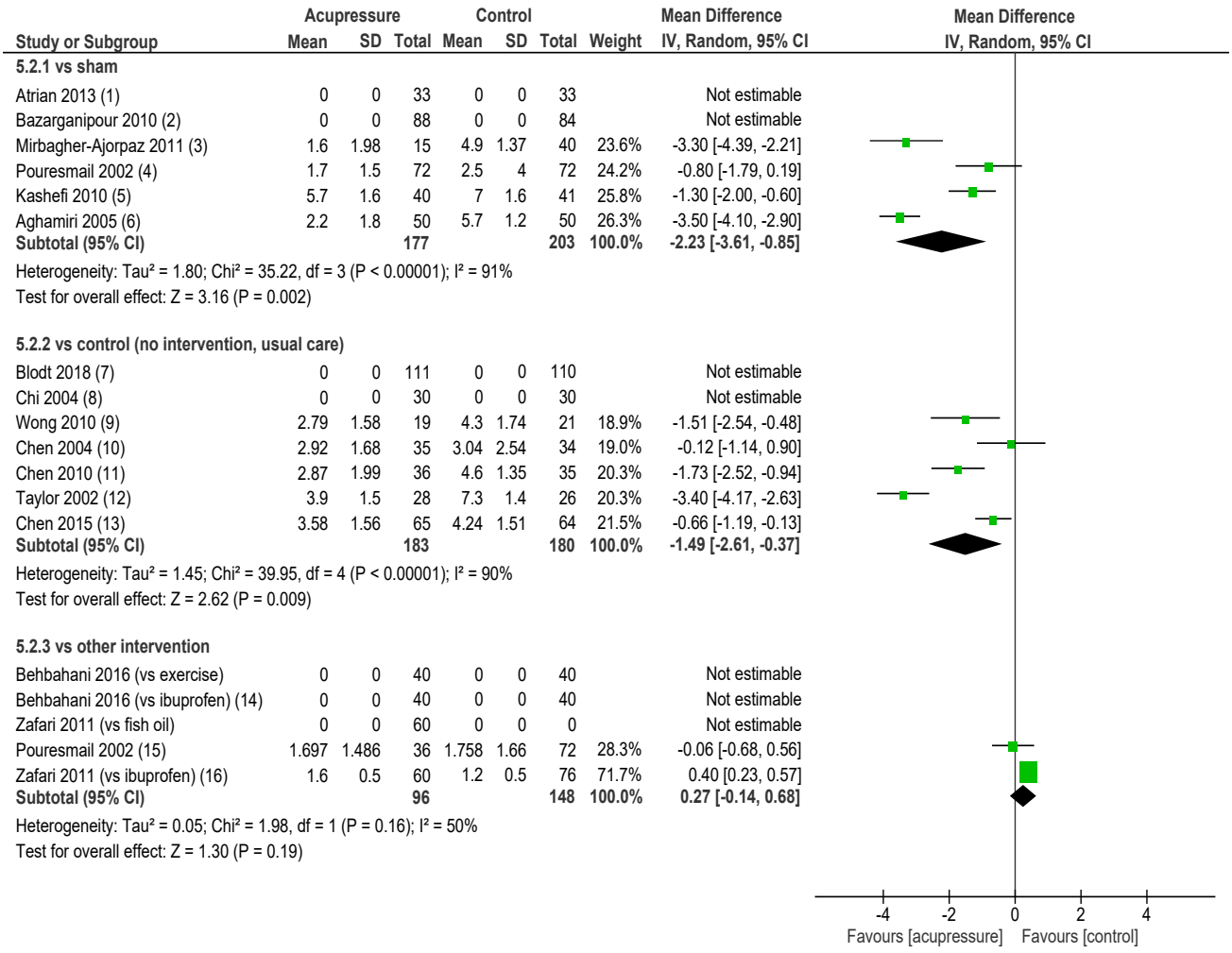
(1) Data from Amour 2019: a greater reduction in symptoms in the acupressure group. No data reported.

(2) Data from Smith 2016: High risk of bias. 3 cycles.

(3) Data from Smith 2016: High risk of bias. 2 cycles

(4) Data from Smith 2016: High risk of bias. 3 cycles

Figure S8 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): dysmenorrhoea – pain

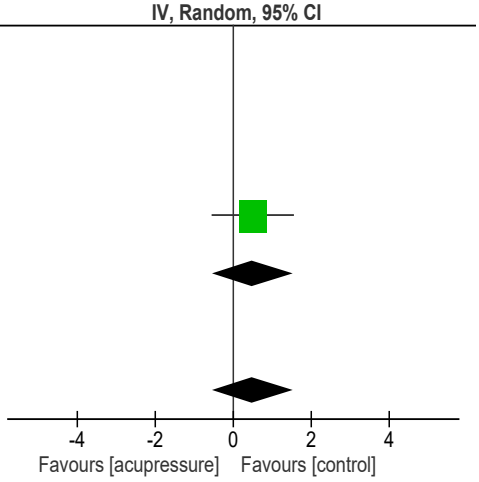


Footnotes

(1) Reported by Tan 2015. 3 cycles. No difference was found between true and sham acupressure groups.
(2) Data from Smith 2016. 2 cycles. Pain severity was lower in the acupressure group than in the control group ($p < 0.001$)
(3) Data from Smith 2016. 1 cycle. Data are skewed.
(4) Data from Kannan 2014. 1 cycle.
(5) Data from Smith 2016. 2 cycles
(6) Data from Smith 2016. High risk of bias. 2 cycles.
(7) Data from Harvie 2019: MD -1.4, 95% CI -2.0, -0.8 reached clinical significance after 6 cycles. Data not reported.
(8) Data from Chung 2012: MD -0.475, 95% CI -0.988 to 0.038. High risk of bias.
(9) Data from Abaraogu 2016. High risk of bias. 3 cycles
(10) Data from Smith 2016: High risk of bias. 2 cycles
(11) Data from Smith 2016: High risk of bias. 3 cycles
(12) Data from Jiang 2013. High risk of bias. 2 cycles
(13) Data from Li 2021. 12 months
(14) Reported by Armour 2019. Hedges g -1.369 (95% CI -1.852, -0.882)
(15) Data from Jiang 2013. vs. Ibuprofen.1 cycle.
(16) Data from Smith 2016. 3 cycles

Figure S9 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): dysmenorrhoea – anxiety

Study or Subgroup	Acupressure Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
5.3.1 vs sham								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
5.3.2 vs control (no intervention, usual care)								
Chen 2004 (1)	3.26	2.23	35	2.76	2.16	34	100.0%	0.50 [-0.54, 1.54]
Chen 2010 (2)	0	0	36	0	0	35		Not estimable
Subtotal (95% CI)			**35**			**34**	**100.0%**	**0.50 [-0.54, 1.54]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.95 (P = 0.34)								
Total (95% CI)			**35**			**34**	**100.0%**	**0.50 [-0.54, 1.54]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.95 (P = 0.34)								
Test for subgroup differences: Not applicable								



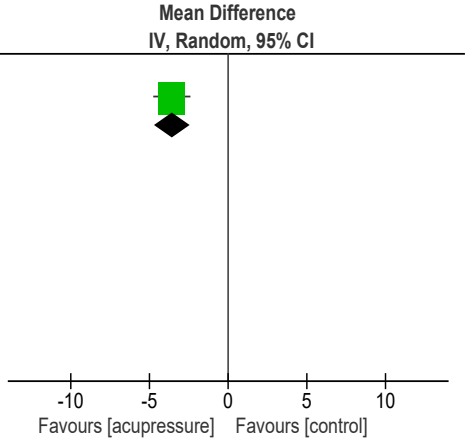
Footnotes

(1) Data from Abaraogu 2016: 2 cycles

(2) 3 cycles.

Figure S10 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): dysmenorrhoea – emotional functioning

Study or Subgroup	Acupressure Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
5.4.1 vs sham								
Kashefi 2010 (1)	3.99	2.69	43	7.57	2.67	43	100.0%	-3.58 [-4.71, -2.45]
Subtotal (95% CI)			**43**			**43**	**100.0%**	**-3.58 [-4.71, -2.45]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.19 (P < 0.00001)								
5.4.2 vs control (no intervention, usual care)								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								



Footnotes

(1) Data from Abaraogu 2016. 2 cycles

S1.11 Pregnancy and childbirth

S1.11.1 Description of studies

Seven (7) citations (81-86) corresponding to 7 systematic reviews (Direkvand-Moghadam 2013, Mollart 2015, Makvandi 2016, Smith 2017, Najafi 2018, Harvie 2019, Smith 2020) were identified in the literature that assessed acupressure compared to sham, control or an active intervention in pregnant females (more than 34 gestational weeks, requiring labour induction or in active labour)[k]. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There were 4 systematic reviews awaiting classification (87-91) (see Appendix C3.2) and no ongoing reviews (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S17. Review details, including all outcome domains and measures and the risk of bias of the included studies are provided in Appendix F1.2.

The RCTs included by the systematic review authors were conducted in pregnant females, typically at-term (37+ gestational weeks) and undergoing spontaneous or induced labour. The studies were judged to be somewhat applicable to the population evaluated in shiatsu (post-term mothers), noting that post-term births are not common in Australia (0.6% of live births) and are associated with worse outcomes (92). Sixteen (16) studies compared acupressure with a sham intervention (Lee 2004, Kashanian 2010, Hjelmstedt 2010, Kordi 2010, Samadi 2010, Salehian 2010, Kordi 2011, Salehian 2011, Aghdam 2012, Hamidzadeh 2012, Dabiri 2013, Sehhatie-Shafaie 2013, Gregson 2015, Akbarzadeh 2015, Mefetoni 2015, Torkzahrani 2016); 18 studies compared acupressure with control (no intervention, usual care) (Chung 2003, Ingram 2005, Hjelmstedt 2010, Kordi 2010, Samadi 2010, Kordi 2011, Salehian 2011, Dabiri 2013, El Hamid 2013, Akbarzadeh 2014, Calik 2014, Mefetoni 2015, Torkzahrani 2015, Mollart 2016, Ozgoli 2016, Torkzahrani 2016, Hamlaci 2017 Mansouri 2018), noting many studies included 2 intervention or comparator groups. One study (Chung 2003) compared acupressure with another intervention (effleurage). The comparator group for 4 RCTs was not recorded by the systematic review authors (Chang 2004, Heidari 2008 Hamidzadeh 2010, Akbarzadeh 2013).

S1.11.2 Critical appraisal

Three reviews (Makvandi 2016, Smith 2017, Smith 2020) were judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest (i.e. met, or partially met, critical AMSTAR-2 domains (4, 8, 9 and 11).

Three (3) systematic reviews (Mollart 2015, Harvie 2019, Najafi 2018) had at least one critical flaw (i.e. did not meet, or partially meet one critical AMSTAR-2 domain [4, 8, 9 or 11]. Of these, one systematic review (Najafi 2018) did not justify publication restrictions, failed to adequately describe the included studies in detail and did not use a satisfactory technique for assessing the risk of bias of individual studies (domains 4, 8 & 9). The other 2 systematic reviews (Mollart 2015, Harvie 2019) did not perform a meta-analysis (domain 11).

One systematic review (Direkvand-Moghadam 2013) did not meet AMSTAR-2 domains 4, 8, 9 or 11.

A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

[k] Systematic reviews that focused on acupressure in early pregnancy or only on labour pain were not prioritised as the evidence in shiatsu was focused on labour induction (see Appendix C5).

Table S17 List of included systematic reviews and overlap with eligible RCTs (per outcome): Pregnancy and childbirth

Review ID	Best available*	SR Outcome domains (measures)	Study ID: Chung 2003	Lee 2004	Chang 2004	Ingram 2005	Heidari 2008	Kashanian 2010	Hjelmstedt 2010	Hamidzadeh 2010	Kordi 2010	Samadi 2010	Salehian 2010	Salehian 2011	Kordi 2011	Aghdam 2012	Hamidzadeh 2012	Akbarzadeh 2013	Dabiri 2013	El Hamid 2013	Sehhatie-Shafaie 2013	Akbarzadeh 2014	Calik 2014	Akbarzadeh 2015	Gregson 2015	Mefetoni 2015	Torkzahrani 2015	Mollart 2016	Ozgoli 2016	Torkzahrani 2016	Hamlaci 2017	Mansouri 2018
Direkvand-Moghadam 2013 (81)	X	Birth experience (bishop score)	--	?	?	--	--	?	--	--	--	--	--	--	--	--	?	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Mollart 2015 (82)	†	Birth experience (labour duration)	?	?	--	?	--	?	?	--	--	--	--	--	--	--	?	--	--	?	--	--	--	--	--	--	--	--	--	--	--	--
		Pregnancy-related pain (VAS)	?	?	--	?	--	?	?	--	--	--	--	--	--	--	?	--	--	?	--	--	--	--	--	--	--	--	--	--	--	--
Makvandi 2016 (83)	✓	Birth experience (labour duration)	Y	Y	--	--	--	Y	!	--	--	!	Y	Y	--	Y	Y	--	?	Y	--	!	?	Y	--	Y	--	--	--	--	--	--
Smith 2017 (93)	✓	Birth experience (bishop score)	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	!	--	Y	!	--	!	--	--
Najafi 2018 (85)	†	Pregnancy-related pain (VAS)	--	Y	--	--	Y	Y	--	Y	Y	Y	Y	Y	Y	--	Y	Y	Y	Y	--	--	--	--	--	Y	--	--	--	--	--	--
Harvie 2019 (11)	†	Birth experience (bishop score)	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	?	!	--	--	--	--
Smith 2020 (86)	✓	Pregnancy-related pain (VAS)	Y	Y	--	--	--	Y	Y	--	Y	--	--	Y	--	--	Y	--	Y	--	Y	--	Y	--	--	Y	--	--	Y	!	Y	Y

Abbreviations: VAS, visual analogue scale

* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 8, domain 9 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])

✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)

Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]

? RCT is included in the systematic review, meets our PICO criteria but a study result is not available [data is incomplete; result may be available in another SR]

! RCT is included in the systematic review, meets our PICO criteria but the SR indicates that the study does not measure the listed outcome [not measured]

-- RCT is not included in the systematic review

S1.11.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in pregnant women are listed in Table S18.

Table S18 Outcomes considered by the NTWC to be critical or important for decision making: Pregnancy and childbirth

Prioritised outcome domain	Measured with	consensus rating	Results available for comparison 1 or 2?	Review ID						
				Direkvand-Moghadam 2013	Mollart 2015	Makvandi 2016	Smith 2017	Najafi 2018	Harvie 2019	Smith 2020
Birth experience	Duration of labour	Critical	Yes	--	†	✓	X	✓	X	✓
Pregnancy-related pain	VAS	Critical	Yes	?	X	?	?	X	?	✓
Quality of life	NR	Critical	No	--	?	?	?	?	?	?
Functional capacity	NR	Critical	No	?	?	?	?	?	?	?
Perceived stress	NR	Critical	No	?	?	?	?	?	?	?
Maternal morbidity	NR	Important	No	?	?	?	?	?	?	--
Foetal health	Apgar score	Important	No	?	?	?	--	?	?	--

Abbreviations: NR, not reported; VAS, visual analogue scale

✓ A study result is available for inclusion in the synthesis

X A study result is available for inclusion, but the systematic review reports incomplete data. Due to time and resource constraints, only the information presented in the systematic review is reported.

† A study result is available for inclusion, but the systematic review only reports the direct of effect. Due to time and resource constraints, only the information presented in the systematic review is reported.

-- The systematic review assessed this outcome but did not find or include any eligible primary studies that reported the outcome.

? The systematic review did not assess this outcome. It is unclear if the outcome was assessed by the primary studies included in the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

The systematic reviews identified 16 RCTs comparing acupressure with sham in pregnant females (requiring labour induction or in labour) that were eligible for this comparison and contributed data to 2 out of 7 critical or important outcomes.

Birth experience (duration of labour)

Several systematic reviews (Mollart 2015, Makvandi, Najafi 2018, Smith 2020) identified 9 RCTs that reported the duration of labour based on a total duration (minutes) or based on time in active stages of labour (stage 1 and/or stage 2) (Lee 2004, Hamidzadeh 2010, Kashanian 2010, Salehian 2010, Aghdam 2012, Hamidzadeh 2012, Dabiri 2013, Akbarzadeh 2015, Mefetoni 2015). It was not always clear when measurements started, but typically was at the onset of regular contractions and when the cervix had dilated between 3 and 7 centimetres. Active labour usually lasts between 240 and 480 minutes.

Pooled results of 6 RCTs (total 559 participants) suggest an effect in favour of acupressure compared with the sham groups (MD –78.82; 95% CI –116.42, –41.23; $p < 0.0001$) but there was substantial statistical heterogeneity ($I^2 = 88\%$; $p < 0.00001$) therefore the standard mean difference was considered (SMD –0.87; 95% CI –1.10, –0.64; $p < 0.00001$; $I^2 = 42\%$).

Data from 3 RCTs were not able to be included in the meta-analysis. Of these, 2 RCTs suggested an effect favouring acupressure in the first stage of labour, but not the second stage, and one RCT suggested there was no statistically significant difference in the duration of the first or second stage of labour between groups.

Labour pain

Several systematic reviews (Mollart 2015, Makvandi, Najafi 2018, Smith 2020) identified 12 RCTs that reported labour pain measured with a 10 cm visual analogue scale or a 0 to 10 numeric pain rating scale (NPRS) (Lee 2004, Heidari 2008, Hjelmstedt 2010, Kashanian 2010, Kordi 2010, Salehian 2010, Samadi 2010, Kordi 2011, Hamidzadeh 2012, Dabiri 2013, Sehhatie-Shafaie 2013, Mefetoni 2015). It was not clear when pain measurements were taken (e.g. end of contractions, end of active labour).

The VAS is a subjective assessment of pain, reported by participants and measured on a continuous scale (cm) from 0 (no pain) to 10 (worst imaginable pain). Higher values indicate worse pain. The NPRS is a segmented numeric version of the visual analogue scale. An MCID for the VAS is not established for labour pain, noting results in this population can be unreliable, possibly due to a 'ceiling effect' (94, 95). The MCID can vary from 8 to 40 mm across different patient groups (mean 17 mm; 95% CI 15 to 19 mm) (60).

Pooled results of 9 RCTs (total 935 participants) suggest an effect in favour of acupressure compared with the sham groups (MD –1.44; 95% CI –2.33, –0.55; $p = 0.002$) but there was substantial statistical heterogeneity ($I^2 = 95\%$; $p < 0.00001$) that did not materially change when the standard mean difference was considered ($I^2 = 89\%$). The clinical importance of the change is not clear.

Data from 3 RCTs (Heidari 2008, Kordi 2010, Kordi 2011) were not able to be included in the meta-analysis (not reported).

Comparison 2 (vs control)

The systematic reviews identified 18 RCTs comparing acupressure with control (usual care) in pregnant females (requiring labour induction or in labour) that were eligible for this comparison, of which 14 RCTs contributed data relevant to 2 of the 7 critical or important outcomes. There were 4 RCTs (Hjelmstedt 2010, Torkzahrani 2015, Mollart 2016, Torkzahrani 2016) that did not contribute any data as they did not report any critical or important outcomes.

Birth experience (duration of labour)

Several systematic reviews (Mollart 2015, Makvandi, Najafi 2018, Smith 2020) identified 9 RCTs that reported the duration of labour based on a total duration (minutes) or based on time in active stages of labour (stage 1 and/or stage 2) (Chung 2003, Ingram 2005, Salehian 2011, Akbarzadeh 2013, Dabiri 2013, El Hamid 2013, Calik 2014, Mefetoni 2015, Hamlaci 2017). It was not always clear when measurements started, but typically was at the onset of regular contractions and when the cervix had dilated between 3 and 7 centimetres. Active labour usually lasts between 240 and 480 minutes.

Pooled results from 4 RCTs (total 338 participants) suggest an effect in favour of acupressure compared with control groups (MD –45.02; 95% CI –76.35, –13.69; $p = 0.005$) but statistical heterogeneity was high ($I^2 = 74\%$; $p < 0.00001$) therefore the standard mean difference was considered (SMD –0.68; 95% CI –0.94, –0.41; $p < 0.00001$; $I^2 = 25\%$).

Data from 5 RCTs were not able to be included in the meta-analysis. Of these, 2 RCTs the data suggested an effect favouring acupressure (El Hamid 2013, Chung 2003), but for 3 RCTs (Ingram 2005, Dabiri 2013, Calik 2014) any observed effect was not statistically significant (data not reported).

Labour pain

Several systematic reviews (Mollart 2015, Makvandi, Najafi 2018, Smith 2020) identified 9 RCTs that reported labour pain measured with a 10 cm visual analogue scale or a 0 to 10 numeric pain rating scale (NPRS) (Kordi 2010, Akbarzadeh 2013, Dabiri 2013, Calik 2014, Mefetoni 2015, Ozgoli 2016, Hamlaci 2017, Mansouri 2018). It was not clear when pain measurements were taken (e.g. end of contractions, end of active labour).

The VAS is a subjective assessment of pain, reported by participants and measured on a continuous scale (cm) from 0 (no pain) to 10 (worst imaginable pain). Higher values indicate worse pain. The NPRS is a segmented numeric version of the visual analogue scale. An MCID for the VAS is not established for labour pain, noting results in this population can be unreliable, possibly due to a 'ceiling effect' (94, 95). The MCID can vary from 8 to 40 mm across different patient groups (mean 17 mm; 95% CI 15 to 19 mm) (60).

Pooled results of 8 RCTs (total 615 participants) suggest an effect in favour of acupressure compared with the control groups (MD –1.72; 95% CI –2.58, –0.85; $p = 0.0001$) but there was substantial statistical heterogeneity ($I^2 = 88\%$; $p < 0.00001$) that did not materially change when the standard mean difference was considered ($I^2 = 82\%$). The clinical importance of the change score is not clear.

Data from 1 RCT (Calik 2014) was not able to be included in the meta-analysis. An effect favouring acupressure was noted ($p < 0.001$).

Comparison 3 (vs active)

The systematic reviews found one RCT (Chung 2003) comparing acupressure with an active intervention (effleurage) in pregnant females (requiring labour induction or in active labour). There were no data reported for this comparison.

S1.11.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There were 16 RCTs identified by the included systematic reviews that compared acupressure with sham in pregnant females (requiring labour induction or in labour) and contributed data to 2 critical or important outcomes.

Acupressure compared to sham for pregnancy and childbirth

Patient or population: pregnant females (requiring labour induction or in active labour)
Setting: community or hospital
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Birth experience assessed with: Duration of labour (mins) (higher is worse) follow-up: immediately after	The mean duration ranged from **360 minutes** (range 185 to 891.4)	MD **78.82 lower** (116.42 lower to 41.23 lower)	-	559 (6 RCTs) # data from 3 RCTs (350 participants) not included here	⨁⨁◯◯ LOW [a,b,c,d,e]	Acupressure may result in a large reduction in the duration of labour in pregnant females.
Quality of life – not reported	-	-	-	-	-	The effect of acupressure on quality of life in pregnant females is unknown
Pain assessed with: VAS or NPRS (higher is worse) Scale from: 0 to 10 follow-up: unclear	The mean pain score was **7.26** points	MD **1.44 lower** (2.33 lower to 0.55 lower)	-	935 (9 RCTs) ## missing data from 3 RCTs (226 participants)	⨁⨁◯◯ LOW [a,c,e,f,g]	Acupressure may result in a reduction in labour pain in pregnant females. **
Perceived stress – not reported	-	-	-	-	-	The effect of acupressure on perceived stress in pregnant females is unknown
Functional capacity – not reported	-	-	-	-	-	The effect of acupressure on functional in pregnant females is unknown

Acupressure compared to sham for pregnancy and childbirth

Patient or population: pregnant females (requiring labour induction or in active labour)
Setting: community or hospital
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Maternal morbidity – not reported	-	-	-	-	-	The effect of acupressure during pregnancy on maternal morbidity is unknown.
Foetal health – not reported	-	-	-	-	-	The effect of acupressure during pregnancy on fetal health at birth is unknown.

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

^ As a rule of thumb, an SMD of 0.2 is considered a small difference, 0.5 is medium, and 0.8 is large difference (14).
^^ In the absence of an MCID the effect estimate was considered based on three levels: small (MD less than 10% of the scale), moderate (MD between 10% to 20% of the scale) or large (MD more than 20% of the scale

** MCID is unknown^^. Mean MCID of 1.7 cm (range 0.8 to 4.0 cm) reported across different patient groups (60).

2 RCTs suggested an effect favouring acupressure in the first stage of labour, but not the second stage, and one RCT suggested there was no difference between groups.
No data reported.

CI: confidence interval; **MD:** mean difference; **NPRS:** numeric pain rating scale; **VAS:** visual analogue scale

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency. Statistical heterogeneity explained by variances in study design and timing of outcome measurement. Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence is in non-complicated pregnancies undergoing spontaneous or induced labour. Certainty of evidence not downgraded.
d. Serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and moderate important difference). Certainty of evidence downgraded.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.
f. No serious inconsistency. All studies are indicating benefit. Statistical heterogeneity explained by variances in study design. Certainty of evidence not downgraded.
g. Serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and small important difference). Certainty of evidence downgraded.

Comparison 2 (vs control)

There were 14 RCTs identified by the included systematic reviews that compared acupressure with control (no intervention, waitlist or usual care) in pregnant females (requiring labour induction or in labour) and contributed data to 2 outcomes.

Acupressure compared to control (no intervention, waitlist, usual care) for pregnancy and childbirth

Patient or population: pregnant females (requiring labour induction or in active labour)
Setting: community or hospital
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Birth experience assessed with: Duration of labour (mins) (higher is worse) follow-up: immediately after	The mean duration ranged from **241.4 minutes (range 216.6 to 913.1)**	MD **45.02 lower** (76.35 lower to 13.69 lower)	-	338 (4 RCTs) # data from 5 RCTs (526 participants) not included here	⨁⨁◯◯ LOW [a,b,c,d,e]	Acupressure may result in a reduction in the duration of labour in pregnant females.
Quality of life – not reported	-	-	-	-	-	The effect of acupressure on quality of life in pregnant females is unknown
Pain assessed with: VAS or NPRS (higher is worse) Scale from: 0 to 10 follow-up: unclear	The mean pain score was **8.235 points**	MD **1.72 lower** (2.58 lower to 0.85 lower)	-	615 (8 RCTs) ## missing data from 1 RCTs (193 participants)	⨁⨁◯◯ LOW [a,c,e,f,g]	Acupressure may result in a reduction in labour pain in pregnant females.**
Perceived stress – not reported	-	-	-	-	-	The effect of acupressure on perceived stress in pregnant females is unknown
Functional capacity – not reported	-	-	-	-	-	The effect of acupressure on functional in pregnant females is unknown
Maternal morbidity – not reported	-	-	-	-	-	The effect of acupressure during pregnancy on maternal morbidity is unknown.
Foetal health – not reported	-	-	-	-	-	The effect of acupressure during pregnancy on fetal health at birth is unknown.

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

^ As a rule of thumb, an SMD of 0.2 is considered a small difference, 0.5 is medium, and 0.8 is large difference (14).
^^ In the absence of an MCID the effect estimate was considered based on three levels: small (MD less than 10% of the scale), moderate (MD between 10% to 20% of the scale) or large (MD more than 20% of the scale

** MCID is unknown^^. Mean MCID of 1.7 cm (range 0.8 to 4.0 cm) reported across different patient groups (60).
# An effect favouring acupressure suggested in 2 RCTs, no effect observed in 3 RCTs (data not reported).

CI: confidence interval; **MD:** mean difference; **NPRS:** numeric pain rating scale; **VAS:** visual analogue scale

Acupressure compared to control (no intervention, waitlist, usual care) for pregnancy and childbirth

Patient or population: pregnant females (requiring labour induction or in active labour)
Setting: community or hospital
Intervention: acupressure
Comparison: control (no intervention, waitlist, usual care)

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				

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

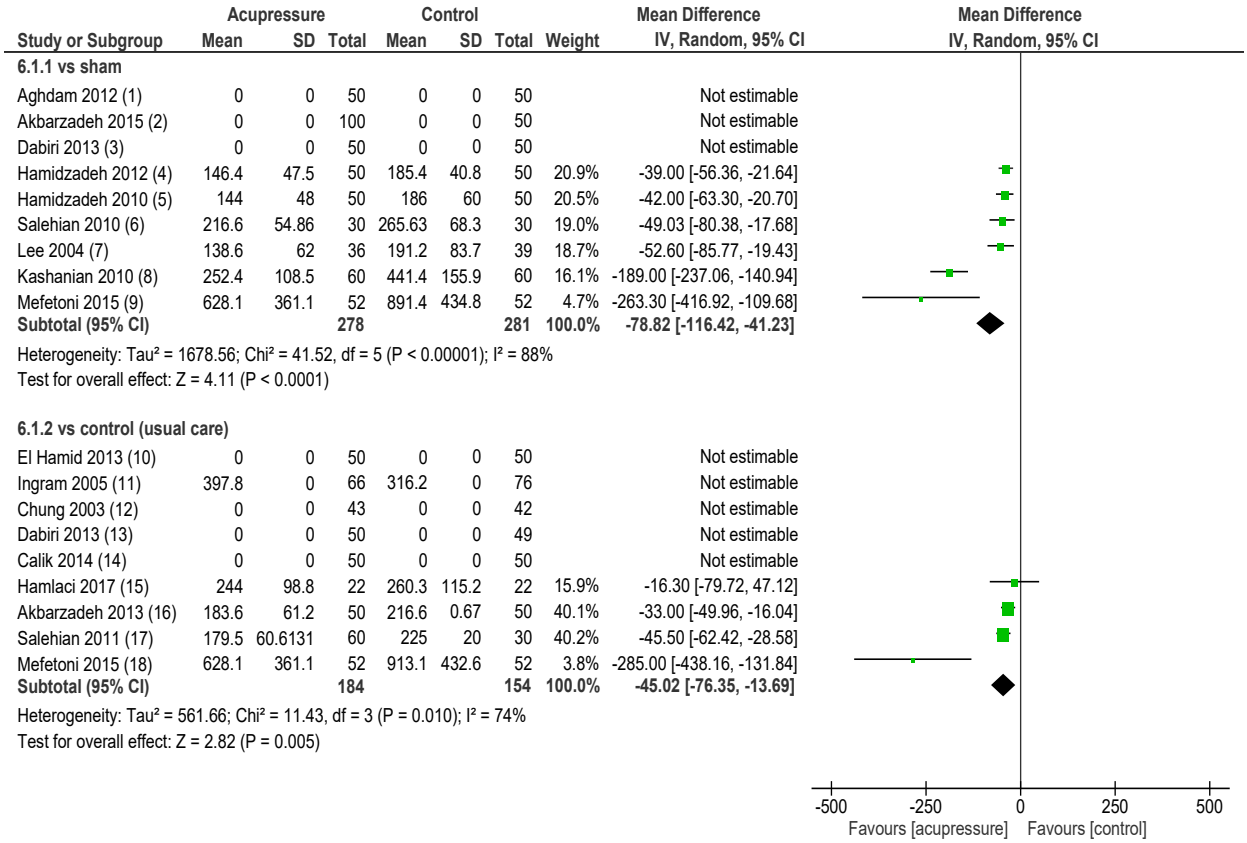
Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. No serious inconsistency. Statistical heterogeneity explained by variances in study design and timing of outcome measurement. Certainty of evidence not downgraded.
c. No serious indirectness. The available evidence in females with non-complicated pregnancies undergoing spontaneous or induced labour. Certainty of evidence not downgraded.
d. Serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and moderate important difference). Certainty of evidence downgraded.
e. Publication bias suspected. There is a strong suspicion of non-reporting of results likely to be related to *p* value, direction or magnitude of effect. Certainty of evidence downgraded.
f. No serious inconsistency. All studies are indicating benefit. Statistical heterogeneity explained by variances in study design. Certainty of evidence not downgraded.
g. Serious imprecision. Wide confidence intervals (upper and lower bounds overlap with both large and small important difference). Certainty of evidence downgraded.

S1.11.5 Forest plots

Outcome results for pregnant females (requiring labour induction or in active labour) (where additional analyses were required and able to be carried out) are presented in Figure S11 (labour duration) and Figure S12 (labour pain).

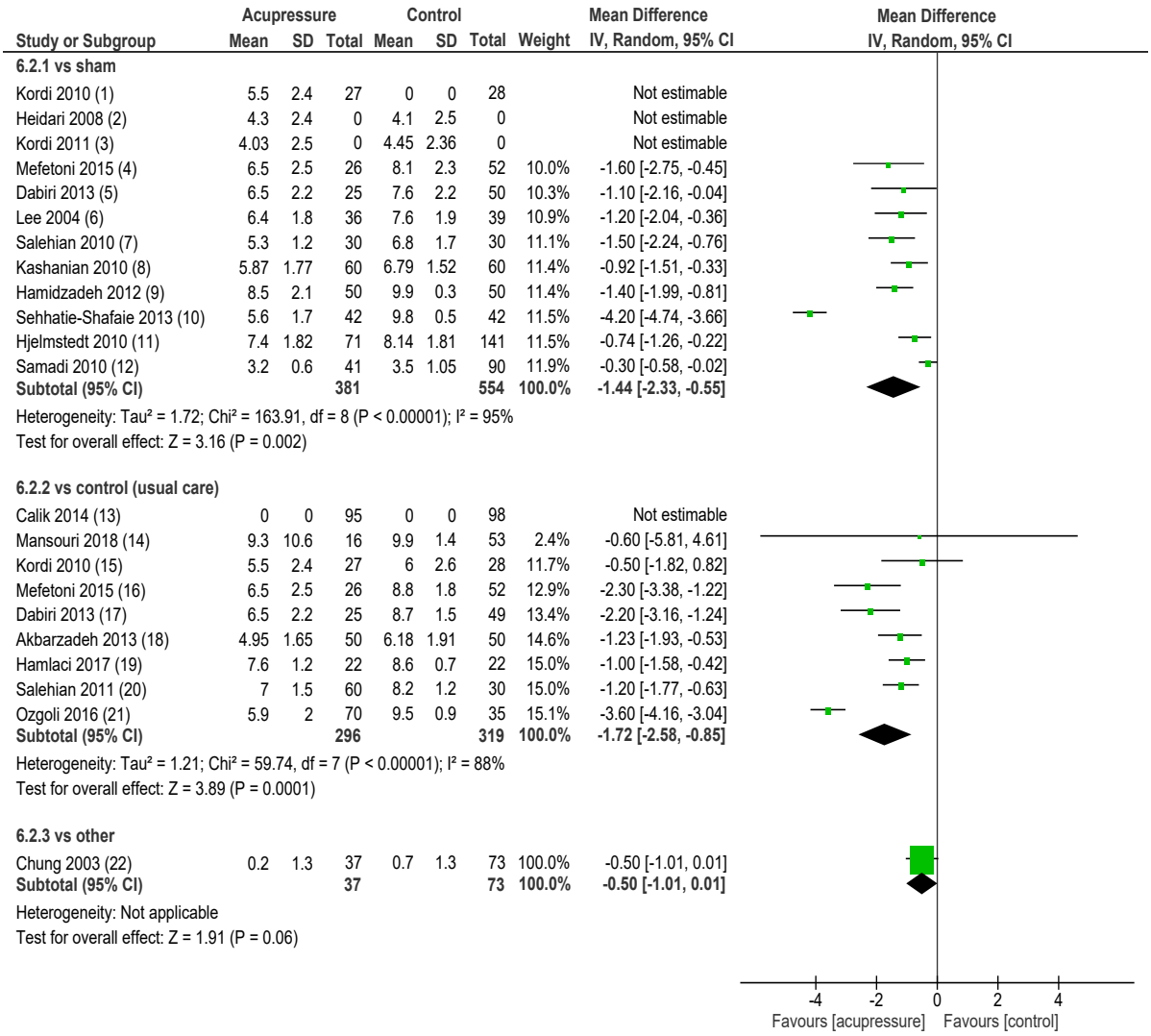
Figure S11 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): pregnancy and childbirth – labour duration (minutes)



Footnotes

(1) Data from Makvandi 2016. 1st stage MD -1.250; 95% CI -1.488 to -1.012; p<0.001; 2nd stage MD -3.4; 95% CI -10.307 to 3.388; p=0.326
(2) Data from Makvandi 2016: Gp1 1st stage MD -0.55; p=0.001; 2nd stage MD -3.72 p=0.268; Gp 2 MD -0.75; p<0.001 & MD -2.99; p=0.400
(3) Data from Makvandi 2016. No statistically significant difference in the duration of the first or second stage of labour between groups.
(4) Data from Smith 2020.
(5) Data from Najafi 2018 (active stage). Comparator group not clear.
(6) Data from Najafi 2018 (active stage).
(7) Data from Smith 2020. Numbers different in Najafi 2018: (active stage) 108.30 mins vs 146.30
(8) Data from Smith 2020.
(9) Data from Smith 2020.
(10) Data from Makvandi 2016. MD -3.430; 95% CI -4.238 to -2.622; p<0.001
(11) Data form Mollart 2015. Shorter (total labour duration) in standard care group; p=0.19. No other data provided.
(12) Data from Makvandi 2016. MD -2.12; 95% CI -3.642 to -0.598; p<0.006
(13) Data from Makvandi 2016. No statistically significant difference in the duration of the first or second stage of labour between groups.
(14) Data from Makvandi 2016. Shorter duration (1st and 2nd stage) in the acupressure group but no data provided.
(15) Data from Smith 2020.
(16) Data from Najafi 2018 (active phase). Comparator group not clear.
(17) Data from Najafi 2018 (active stage). Data for 2 acupressure groups (SP6 & LI4) combined.
(18) Data from Makvandi 2016.

Figure S12 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): pregnancy and childbirth – labour pain



Footnotes

(1) Data from Smith 2020. Comparator group data not available (not in English).
(2) Data from Najafi 2018. Comparator group not known. Group numbers not reported (total 128) (study not in English)
(3) Data from Najafi 2018. Comparator group not clear (sham & usual care). Group numbers not reported (total 102) (study not in English)
(4) Data from Smith 2020.
(5) Data from Smith 2020.
(6) Data from Najafi 2018.
(7) Data form Najafi 2018.
(8) Data from Smith 2020.
(9) Data from Smith 2020. Data from Najafi 2018 does not match (different timepoints)
(10) Data from Smith 2020.
(11) Data from Smith 2020. Control groups are combined (sham + usual care)
(12) Data from Najafi 2018. Not clear of complarator group (assumed sham and control group data combined).
(13) Data from Smith 2020. Data are in medians and not able to be included. An effect favouring acupressure noted (p<0.001)
(14) Data from Smith 2020.
(15) Data from Smith 2020.
(16) Data from Smith 2020.
(17) Data from Smith 2020.
(18) Data from Najafi 2018. Comparator group not clear.
(19) Data from Smith 2020.
(20) Data from Smith 2020. 2 acupressure groups combined. Data from Najafi 2018 does not match (different timepoints?)
(21) Data from Smith 2020.
(22) Data from Smith 2020. COmbined control groups (effleurage + usual care)

S1.12 Recovery after minimally invasive surgery

S1.12.1 Description of studies

Three citations (36, 96, 97) corresponding to 3 systematic reviews (Hewitt 2009, Lee 2015, Waits 2018) were identified in the literature that assessed acupressure compared to sham, control or an active intervention in patients recovering after minimally invasive surgery. No additional reviews were identified in the Departments public call for evidence (see Appendix C2). There are 3 systematic reviews awaiting classification (98-100) (see Appendix C3.2) and one ongoing review (101) (see Appendix C4.2).

An overview of the included systematic reviews and their overlap with eligible RCTs is provided in Table S19. Review details, including all outcome domains and measures and the risk of bias of the included studies are provided in Appendix F1.2.

The studies were in people receiving minimally invasive surgery (gynaecological laparoscopy, endoscopic retrograde cholangiopancreatography) and are directly applicable the populations evaluated in shiatsu (laparoscopic surgery)[l]. Eight studies compared acupressure with sham intervention (Harmon 1999, Schlager 2001, Agarwal 2002, Boehler 2002, Samad 2003, Sadigha 2008, Iqbal 2012, Liu 2012, White 2012) and one study (Liu 2012) compared acupressure with control (no intervention) described as sleep hygiene education.

Table S19 List of included systematic reviews and overlap with eligible RCTs (per outcome): Recovery after minimally invasive surgery

Review ID	Best available*	SR Outcome domains (measures)	Study ID								
			Harmon 1999	Schlager 2001	Agarwal 2002	Boehler 2002	Samad 2003	Sadigha 2008	Iqbal 2012	Liu 2012	White 2012
Hewitt 2009 (96)	✓	Post-operative complaints (nausea and vomiting)	--	Y	--	Y	--	--	--	--	--
Lee 2015 (97)	✓	Post-operative complaints (nausea and vomiting)	Y	--	Y	--	Y	Y	Y	--	Y
Waits 2018 (36)	✓	Sleep quality	--	--	--	--	--	--	--	Y	--

* Best available information (in any order) means the systematic review meets AMSTAR-2 domain 4, domain 8, domain 9 and domain 11 (see Framework for selecting the systematic review from which to extract data [Appendix B2])
✓ Systematic review meets (or partially meets) prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
† Systematic review meets (or partially meets) some, but not all, prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
X Systematic review does not meet prespecified critical AMSTAR-2 domains (4, 8, 9 & 11)
Y RCT is included in the systematic review, meets our PICO criteria & a study result is reported for the listed outcome measure [result available]
? RCT is included in the systematic review, meets our PICO criteria but a study result is not available for the listed outcome measure [data is incomplete; result may be available in another SR]
-- RCT is not included in the systematic review

S1.12.2 Critical appraisal

All 3 included systematic reviews (Hewitt 2009, Lee 2015, Waits 2018) were judged to probably provide an accurate and comprehensive summary of the available studies that address the question of interest (i.e. met, or partially met, critical AMSTAR-2 domains (4, 8, 9 and 11).

[l] There were 11 other systematic reviews that focused on acupressure in people recovering after other types of surgery (e.g. major cardiac, obstetric) that are not included here as evidence in shiatsu was focused on minimally invasive procedures. (see Appendix C5).

A summary of the strengths or limitations of the included systematic reviews assessed against each AMSTAR-2 domain is provided in Appendix E2.

S1.12.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision making in people recovering after minimally invasive surgery are listed in Table S20.

Table S20 Outcomes considered by the NTWC to be critical or important for decision making: Recovery after minimally invasive surgery

Prioritised outcome domain	Measured with	Consensus rating	Results available for comparison 1 or 2?	Review ID		
				Hewitt 2009	Lee 2015	Waits 2018
Clinical recovery	No measures reported in eligible studies	Critical	No	?	?	?
Post-operative complaints	Incidence of nausea and vomiting	Critical	Yes	✓	✓	?
Post-operative pain	Visual analogue scale	Critical	No	?	?	?
Bowel recovery	Time between first and last defecation	Important	No	?	?	?
Pulmonary function	Oxygen saturation	Critical	No	?	?	?

Abbreviations: VAS, visual analogue scale

✓ A study result is available for inclusion in the synthesis

? The systematic review did not assess this outcome. It is unclear if the outcome was assessed by the primary studies included in the SR. Due to time and resource constraints, only the information presented in the systematic review is reported.

Comparison 1 (vs sham)

The systematic reviews (Hewitt 2009, Lee 2015) identified 8 RCTs comparing acupressure with sham in people recovering after minimally invasive surgery that were eligible for this comparison and contributed data to one out of 5 critical or important outcomes.

Post-operative complaints

Two systematic reviews (Hewitt 2009, Lee 2015) included 8 RCTs (Harmon 1999, Schlager 2001, Agarwal 2002, Boehler 2002, Samad 2003, Sadigha 2008, Iqbal 2012, White 2012) that reported the incidence of both nausea and vomiting 0 to 24 hours post-operative.

Pooled results (total 606 participants) suggest the number of nausea episodes was reduced in the acupressure group (121/305) compared with the sham group (168/301) (RR 0.71; 95% CI 0.52, 0.98; $p = 0.04$; $I^2 = 68\%$). Similarly, the number of vomiting episodes was reduced in the acupressure group (49/308) compared with the sham group (127/289) (RR 0.37; 95% CI 0.20, 0.68; $p = 0.001$; $I^2 = 74\%$).

Comparison 2 (vs control)

There were no RCTs found by the included systematic reviews comparing acupressure with control (no intervention, usual care) in people recovering after minimally invasive surgery that contributed data relevant to the 5 critical or important outcomes.

Comparison 3 (vs active)

There were no RCTs found by the included systematic reviews comparing acupressure with an active comparator in people recovering after minimally invasive surgery.

S1.12.4 Summary of findings and evidence statements

Comparison 1 (vs sham)

There were 8 RCTs found by the included systematic reviews comparing acupressure with sham in people recovering after minimally invasive surgery that contributed data to 2 critical or important outcomes.

Acupressure compared to sham for recovery after minimally invasive surgery

Patient or population: recovery after minimally invasive surgery
Setting: hospital
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				
Clinical recovery – not reported	-	-	-	(0 studies)	-	The effect of acupressure on clinical recovery in people recovering after minimally invasive surgery is unknown.
Post-operative nausea – total episodes (higher is worse) Follow-up: 0 to 24 hours	558 per 1000	**396 per 1000** (290 to 547)	**RR 0.71** (0.52 to 0.98) ^	606 (8 RCTs)	⊕⊕◯◯ LOW [a,b,c,d,e]	Acupressure may result in a slight reduction in post-operative nausea in people recovering after minimally invasive (laparoscopic) surgery.
Post-operative pain – not reported	-	-	-	(0 studies)	-	The effect of acupressure on post-operative pain in people recovering after minimally invasive surgery is unknown.
Bowel recovery – not reported	-	-	-	(0 studies)	-	The effect of acupressure on bowel recovery in people recovering after minimally invasive surgery is unknown.
Pulmonary function – not reported	-	-	-	(0 studies)	-	The effect of acupressure on pulmonary function in people recovering after minimally invasive surgery is unknown.
Post-operative vomiting – total episodes (higher is worse) Follow-up: 0 to 24 hours	439 per 1000	**163 per 1000** (88 to 299)	**RR 0.37** (0.20 to 0.68) ^	597 (8 RCTs)	⊕⊕⊕◯ MODERATE [a,b,c,e,f]	Acupressure probably results in a large reduction in post-operative vomiting in people recovering after minimally invasive (laparoscopic) surgery.

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

^ A 25% relative reduction was considered important (i.e. RR < 0.75).

CI: confidence interval; **MD:** mean difference; **NPRS:** numeric pain rating scale; **VAS:** visual analogue scale

Acupressure compared to sham for recovery after minimally invasive surgery

Patient or population: recovery after minimally invasive surgery
Setting: hospital
Intervention: acupressure
Comparison: sham

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	Nº of participants (studies)	Certainty of the evidence (GRADE)	Evidence statement
	Risk with control	Risk with acupressure				

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

Explanations

a. No serious risk of bias. Certainty of evidence not downgraded.
b. Serious inconsistency. Statistical heterogeneity is high ($I^2 > 68\%$) and not able to be explained. Certainty of evidence downgraded.
c. No serious indirectness. The available evidence in people undergoing minimally invasive surgery. It is not clear if the type of acupressure applied is considered part of shiatsu (Korean hand point K-K9 or SP6 bands). Certainty of evidence not downgraded.
d. Serious imprecision. Wide confidence intervals (lower bounds overlap with no important difference). Certainty of evidence downgraded.
e. Publication bias not suspected. Certainty of evidence not downgraded.
f. No serious imprecision. Certainty of evidence not downgraded.

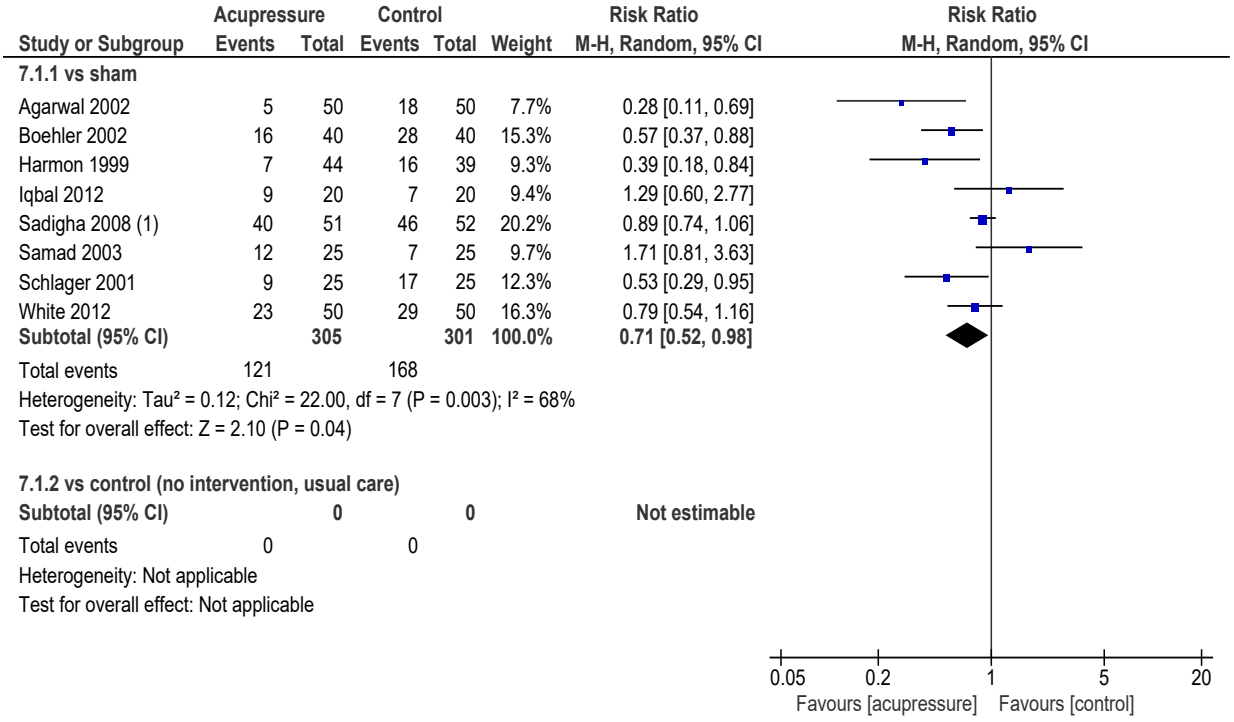
Comparison 2 (vs control)

No studies found.

S1.12.5 Forest plots

Outcome results for recovery after minimally invasive surgery (where additional analyses were required and able to be carried out) are presented in Figure S13 (incidence of nausea) and Figure S14 (incidence of vomiting).

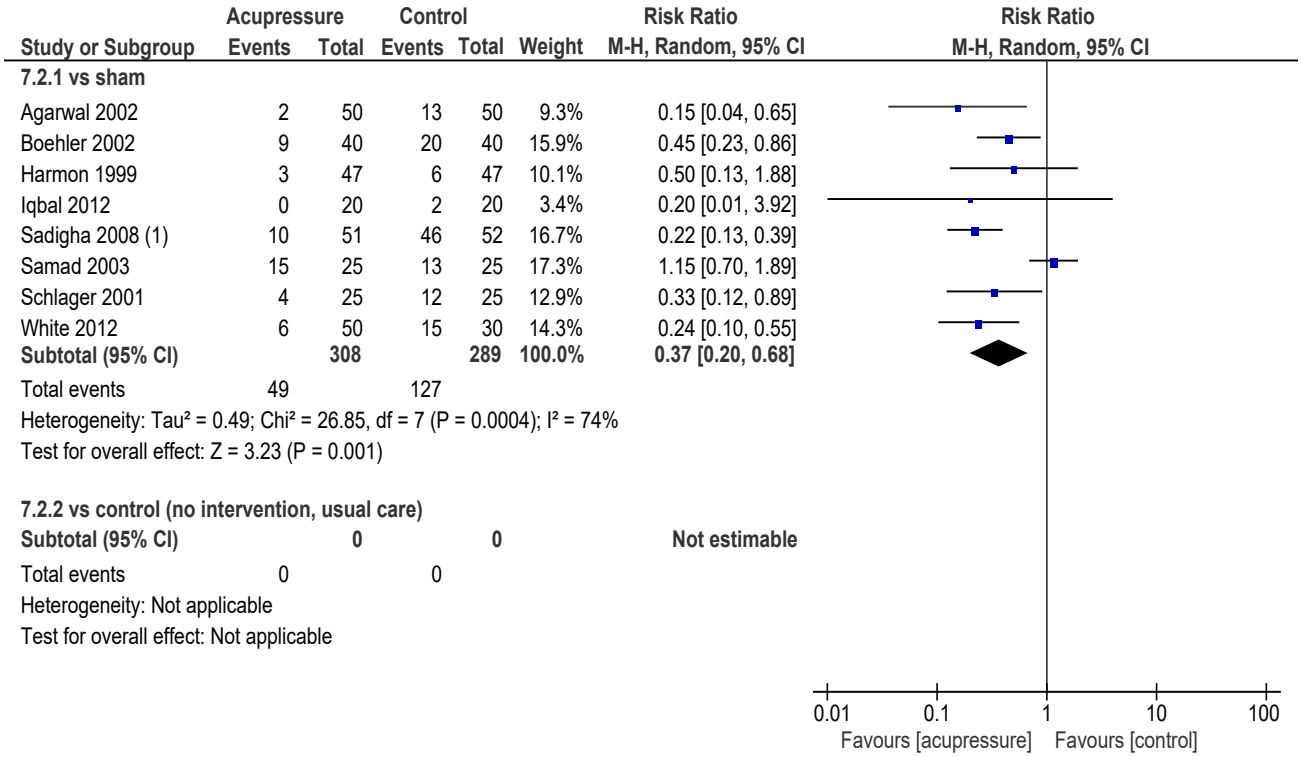
Figure S13 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): recovery after minimally invasive surgery – incidence of nausea (0-24 hours)



Footnotes

(1) High risk of bias.

Figure S14 Forest plot of comparison: Acupressure vs sham or control (no intervention, waitlist, usual activities): recovery after minimally invasive surgery – incidence of vomiting (0-24 hours)



Footnotes

(1) High risk of bias.

Acknowledgments

The Research Protocol was written and developed by **HT**ANALYSTS in conjunction with NHMRC. Expert advice was provided by NTREAP and NTWC, especially in relation to intervention, study design and eligibility criteria. A methodological review of the draft protocol was conducted by Cochrane Australia.

Contributions of authors

The Evidence Evaluation Report was written and developed by **HT**ANALYSTS, with evidence synthesis (statistical analysis and GRADE) conducted by the following reviewers: Margaret Jorgensen (oversight), Aiya Taylor (lead) or Ella Connor. Expert advice was provided by NTREAP and NTWC, especially in relation to intervention, study design and eligibility criteria.

A methodological review of the draft evaluation report was conducted by Cochrane Australia.

Declarations of interest

All named authors declare they have no financial, personal or professional interests that could be construed to have influenced the conduct or results of this systematic review.

In line with the process to establish any NHMRC committee, each committee member was asked to disclose their interests. Potential conflicts of interest among NHMRC NTWC members are lodged with NHMRC and are available online.

REFERENCES

1. Department of Health and Aged Care. The 2015 review of the Australian Government rebate on private health Iinsurance for natural therapies. 2019. [Accessed: 15 December 2019]. Available from: https://www.health.gov.au/internet/main/publishing.nsf/Content/phi-natural-therapies.
2. Higgins JPT, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, et al. Methodological Expectations of Cochrane Intervention Reviews. [Internet]. 2019. [Accessed: October 2019]. Available from: https://community.cochrane.org/mecir-manual.
3. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097. 10.1371/journal.pmed.1000097
4. Churchill R, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, et al. Standards for the reporting of new Cochrane Intervention Reviews. 2019 [cited 4 June 2020]. In: Methodological Expectations of Cochrane Intervention Reviews [Internet]. London: Cochrane, [cited 4 June 2020]. Available from: https://community.cochrane.org/mecir-manual/standards-reporting-new-cochrane-intervention-reviews-r1-r109.
5. Schünemann H, Brożek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach [Internet]. 2013. [Accessed; (Updated October 2013). Available from: https://gdt.gradepro.org/app/handbook/handbook.html.
6. Lee EJ, Frazier SK. The efficacy of acupressure for symptom management: a systematic review. J Pain Symptom Manage. 2011;42(4):589-603. 10.1016/j.jpainsymman.2011.01.007
7. Ling WM, Lui LY, So WK, Chan K. Effects of acupuncture and acupressure on cancer-related fatigue: a systematic review. Oncol Nurs Forum. 2014;41(6):581-92. 10.1188/14.ONF.581-592
8. Duong N, Davis H, Robinson PD, Oberoi S, Cataudella D, Culos-Reed SN, et al. Mind and body practices for fatigue reduction in patients with cancer and hematopoietic stem cell transplant recipients: A systematic review and meta-analysis. Crit Rev Oncol Hematol. 2017;120:210-6. 10.1016/j.critrevonc.2017.11.011
9. Arring NM, Barton DL, Brooks T, Zick SM. Integrative Therapies for Cancer-Related Fatigue. Cancer J. 2019;25(5):349-56. 10.1097/PPO.0000000000000396
10. Calcagni N, Gana K, Quintard B. A systematic review of complementary and alternative medicine in oncology: Psychological and physical effects of manipulative and body-based practices. PLoS ONE. 2019;14(10):e0223564. 10.1371/journal.pone.0223564
11. Harvie A, Steel A, Wardle J. Traditional Chinese Medicine Self-Care and Lifestyle Medicine Outside of Asia: A Systematic Literature Review. J Altern Complement Med. 2019;25(8):789-808. 10.1089/acm.2018.0520
12. Liu XL, Cheng HL, Moss S, Wang CC, Turner C, Tan JY. Somatic Acupoint Stimulation for Cancer-Related Sleep Disturbance: A Systematic Review of Randomized Controlled Trials. Evid Based Complement Alternat Med. 2020;2020:2591320. 10.1155/2020/2591320
13. Hu J, Shen Y, Zhang G, He J, Sun M, Zhang H, et al. Effect of acupoint therapies on chemotherapy-induced nausea and vomiting: A systematic review protocol. Medicine (Baltimore). 2019;98(37):e17109. 10.1097/MD.0000000000017109
14. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. P T. 2008;33(12):700-11.
15. Robinson N, Lorenc A, Liao X. The evidence for Shiatsu: a systematic review of Shiatsu and acupressure. BMC Complement Altern Med. 2011;11:88. 10.1186/1472-6882-11-88
16. Ernst E. Acupuncture/acupressure for weight reduction? A systematic review. Wien Klin Wochenschr. 1997;109(2):60-2.
17. Lee JS, Lee MS, Min K, Lew JH, Lee BJ. Acupressure for treating neurological disorders: a systematic review. Int J Neurosci. 2011;121(8):409-14. 10.3109/00207454.2011.570465
18. Strom BS, Ytrehus S, Grov EK. Sensory stimulation for persons with dementia: a review of the literature. J Clin Nurs. 2016;25(13-14):1805-34. 10.1111/jocn.13169

19. Liu F, Shen C, Yao L, Li Z. Acupoint massage for managing cognitive alterations in older adults: A systematic review and meta-analysis. [References]: The Journal of Alternative and Complementary Medicine. Vol.24(6), 2018, pp. 532-540.; 2018.

20. Hmwe NTT, Browne G, Mollart L, Allanson V, Chan SW. An integrative review of acupressure interventions for older people: A focus on sleep quality, depression, anxiety, and agitation. Int J Geriatr Psychiatry. 2019;34(3):381-96. 10.1002/gps.5031

21. Margenfeld F, Klocke C, Joos S. Manual massage for persons living with dementia: A systematic review and meta-analysis. Int J Nurs Stud. 2019;96:132-42. 10.1016/j.ijnurstu.2018.12.012

22. O'Caoimh R, Mannion H, Sezgin D, O'Donovan MR, Liew A, Molloy DW. Non-pharmacological treatments for sleep disturbance in mild cognitive impairment and dementia: A systematic review and meta-analysis. Maturitas. 2019;127:82-94. 10.1016/j.maturitas.2019.06.007

23. Chen MC, Yang LY, Chen KM, Hsu HF. Systematic Review and Meta-Analysis on Using Acupressure to Promote the Health of Older Adults. J Appl Gerontol. 2020;39(10):1144-52. 10.1177/0733464819870027

24. American Psychological Association. Neuropsychiatric Inventory 2011 [cited 16 August 2022]. Available from: https://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/neuropsychiatric-inventory.

25. Cummings J. Neuropsychiatric Inventory 2009 [cited 16 August 2022]. Available from: https://dementiaresearch.org.au/wp-content/uploads/2016/01/NPI.pdf.

26. Mao HF, Kuo CA, Huang WN, Cummings JL, Hwang TJ. Values of the Minimal Clinically Important Difference for the Neuropsychiatric Inventory Questionnaire in Individuals with Dementia. J Am Geriatr Soc. 2015;63(7):1448-52. 10.1111/jgs.13473

27. Marvin K, Zeltzer L. Barthel Index (BI) 2015 [cited 13 December 2022]. Available from: https://strokengine.ca/en/assessments/barthel-index-bi/#:~:text=The%20Barthel%20Index%20(BI)%20measures,transfer%2C%20ambulation%20and%20stair%20climbing.

28. Andrew MK, Rockwood K. A five-point change in Modified Mini-Mental State Examination was clinically meaningful in community-dwelling elderly people. J Clin Epidemiol. 2008;61(8):827-31. 10.1016/j.jclinepi.2007.10.022

29. Cao H, Pan X, Li H, Liu J. Acupuncture for treatment of insomnia: A systematic review of randomized controlled trials. [References]: The Journal of Alternative and Complementary Medicine. Vol.15(11), 2009, pp. 1171-1186.; 2009.

30. Sarris J, Byrne GJ. A systematic review of insomnia and complementary medicine. [References]: Sleep Medicine Reviews. Vol.15(2), 2011, pp. 99-106.; 2011.

31. Yeung WF, Chung KF, Poon MM, Ho FY, Zhang SP, Zhang ZJ, et al. Acupressure, reflexology, and auricular acupressure for insomnia: a systematic review of randomized controlled trials. Sleep Med. 2012;13(8):971-84. 10.1016/j.sleep.2012.06.003

32. Tan JY, Suen LK, Wang T, Molassiotis A. Sham Acupressure Controls Used in Randomized Controlled Trials: A Systematic Review and Critique. PLoS ONE. 2015;10(7):e0132989. 10.1371/journal.pone.0132989

33. Hmwe NT, Subramaniam P, Tan LP. Effectiveness of Acupressure in Promoting Sleep Quality: A Systematic Review of Randomized Controlled Trials. Holist Nurs Pract. 2016;30(5):283-93. 10.1097/HNP.0000000000000165

34. Wang Z, Hu X, Su J, Gao X, Xu N, Xing Y, et al. The efficacy and safety stimulating a single acu-point shenmen (HT 7) for managing insomnia: A systematic review of randomized controlled trials. European Journal of Integrative Medicine. 2017;15:17-22. 10.1016/j.eujim.2017.08.010

35. Capezuti E, Sagha Zadeh R, Pain K, Basara A, Jiang NZ, Krieger AC. A systematic review of non-pharmacological interventions to improve nighttime sleep among residents of long-term care settings. BMC Geriatr. 2018;18(1):143. 10.1186/s12877-018-0794-3

36. Waits A, Tang YR, Cheng HM, Tai CJ, Chien LY. Acupressure effect on sleep quality: A systematic review and meta-analysis. Sleep Med Rev. 2018;37:24-34. 10.1016/j.smrv.2016.12.004

37. Shang B, Yin H, Jia Y, Zhao J, Meng X, Chen L, et al. Nonpharmacological interventions to improve sleep in nursing home residents: A systematic review. Geriatr Nurs. 2019;40(4):405-16. 10.1016/j.gerinurse.2019.01.001

38. Samara MT, Huhn M, Chiocchia V, Schneider-Thoma J, Wiegand M, Salanti G, et al. Efficacy, acceptability, and tolerability of all available treatments for insomnia in the elderly: a systematic review and network meta-analysis. *Acta Psychiatr Scand*. 2020;142(1):6-17. 10.1111/acps.13201

39. Ames J. Acupressure to treat insomnia: a systematic review. *Sleep Medicine*. 2019;64:S14. 10.1016/j.sleep.2019.11.042

40. Ji XW, Ivers H, Beaulieu-Bonneau S, Morin CM. Complementary and alternative treatments for insomnia/sleep depression-anxiety symptom cluster: meta-analysis of english and Chinese literature. *Sleep Medicine*. 2019;64:S177. 10.1016/j.sleep.2019.11.488

41. Wu DJ, Dong HC, Tang TN, Zhu SF. Acupressure for insomnia: A protocol for systematic review and meta-analysis. *Medicine (Baltimore)*. 2018;97(45):e13180. 10.1097/MD.0000000000013180

42. Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193-213. 10.1016/0165-1781(89)90047-4

43. Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. *Sleep Med Rev*. 2016;25:52-73. 10.1016/j.smrv.2015.01.009

44. Passos MH, Silva HA, Pitangui AC, Oliveira VM, Lima AS, Araujo RC. Reliability and validity of the Brazilian version of the Pittsburgh Sleep Quality Index in adolescents. *J Pediatr (Rio J)*. 2017;93(2):200-6. 10.1016/j.jped.2016.06.006

45. Longo UG, Berton A, De Salvatore S, Piergentili I, Casciani E, Faldetta A, et al. Minimal Clinically Important Difference and Patient Acceptable Symptom State for the Pittsburgh Sleep Quality Index in Patients Who Underwent Rotator Cuff Tear Repair. *Int J Environ Res Public Health*. 2021;18(16). 10.3390/ijerph18168666

46. Quality Metric Incorporated, editor. User's manual for the SF-36v2 survey. 3rd ed2011.

47. Sterling M. General Health Questionnaire - 28 (GHQ-28). *J Physiother*. 2011;57(4):259. 10.1016/S1836-9553(11)70060-1

48. Sibbritt D, Peng W, Lauche R, Ferguson C, Frawley J, Adams J. Efficacy of acupuncture for lifestyle risk factors for stroke: A systematic review. [References]: PLoS ONE. Vol.13(10), 2018, ArtID e0206288.; 2018.

49. Academy of the American Physical Therapy Association. Adult vital sign interpretation in acute care guide. 2021. [Accessed. Available from: https://cardiopt.memberclicks.net/assets/docs/CPG/Joint%20Vital%20Sign%20Booklet.pdf.

50. Ducrotte P, Causse C. The Bowel Function Index: a new validated scale for assessing opioid-induced constipation. *Curr Med Res Opin*. 2012;28(3):457-66. 10.1185/03007995.2012.657301

51. Nikjooy AP, Jafari HP, Saba MAMP, Ebrahimi NMP, Mirzaei RM. Patient Assessment of Constipation Quality of Life Questionnaire: Translation, Cultural Adaptation, Reliability, and Validity of the Persian Version. *Iran J Med Sci*. 2018;43(3):261-8.

52. Soreca V. Quality of Life Measures in Opioid-Induced Constipation: What Constitutes a Minimal Clinically Important Difference? 2021 [cited 16 February 2022]. Available from: https://www.gastroenterologyadvisor.com/general-gastroenterology/quality-of-life-measures-in-opioid-induced-constipation-what-constitutes-a-minimal-clinically-important-difference/.

53. Kim YC, Lee MS, Park E-S, Lew J-H, Lee B-J. Acupressure for the Treatment of Musculoskeletal Pain Conditions: A Systematic Review. *Journal of Musculoskeletal Pain*. 2012;20(2):116-21. 10.3109/10582452.2012.673543

54. Chen YW, Wang HH. The effectiveness of acupressure on relieving pain: a systematic review. *Pain Manag Nurs*. 2014;15(2):539-50. 10.1016/j.pmn.2012.12.005

55. Yuan QL, Guo TM, Liu L, Sun F, Zhang YG. Traditional Chinese medicine for neck pain and low back pain: a systematic review and meta-analysis. *PLoS ONE*. 2015;10(2):e0117146. 10.1371/journal.pone.0117146

56. Yeganeh M, Baradaran HR, Qorbani M, Moradi Y, Dastgiri S. The effectiveness of acupuncture, acupressure and chiropractic interventions on treatment of chronic nonspecific low back pain in Iran: A systematic review and meta-analysis. *Complement Ther Clin Pract*. 2017;27:11-8. 10.1016/j.ctcp.2016.11.012

57. Godley E, Smith MA. Efficacy of acupressure for chronic low back pain: A systematic review. *Complement Ther Clin Pract*. 2020;39:101146. 10.1016/j.ctcp.2020.101146

58. Li T, Li X, Huang F, Tian Q, Fan ZY, Wu S. Clinical Efficacy and Safety of Acupressure on Low Back Pain: A Systematic Review and Meta-Analysis. *Evid Based Complement Alternat Med*. 2021;2021:8862399. 10.1155/2021/8862399

59. Farrar JT, Young JP, Jr., LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94(2):149-58. 10.1016/S0304-3959(01)00349-9

60. Olsen MF, Bjerre E, Hansen MD, Hilden J, Landler NE, Tendal B, et al. Pain relief that matters to patients: systematic review of empirical studies assessing the minimum clinically important difference in acute pain. *BMC Med*. 2017;15(1):35. 10.1186/s12916-016-0775-3

61. Mease PJ, Spaeth M, Clauw DJ, Arnold LM, Bradley LA, Russell IJ, et al. Estimation of minimum clinically important difference for pain in fibromyalgia. *Arthritis Care Res (Hoboken)*. 2011;63(6):821-6. 10.1002/acr.20449

62. Jordan K, Dunn KM, Lewis M, Croft P. A minimal clinically important difference was derived for the Roland-Morris Disability Questionnaire for low back pain. *J Clin Epidemiol*. 2006;59(1):45-52. 10.1016/j.jclinepi.2005.03.018

63. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)*. 2008;33(1):90-4. 10.1097/BRS.0b013e31815e3a10

64. Stratford PW, Riddle DL. A Roland Morris Disability Questionnaire Target Value to Distinguish between Functional and Dysfunctional States in People with Low Back Pain. *Physiother Can*. 2016;68(1):29-35. 10.3138/ptc.2014-85

65. Chang YJ, Lee JS, Lee CG, Lee WS, Lee KS, Bang SM, et al. Assessment of clinical relevant fatigue level in cancer. *Support Care Cancer*. 2007;15(7):891-6. 10.1007/s00520-007-0219-x

66. White AR. A review of controlled trials of acupuncture for women's reproductive health care. *J Fam Plann Reprod Health Care*. 2003;29(4):233-6. 10.1783/147118903101197863

67. Cho SH, Hwang EW. Acupressure for primary dysmenorrhoea: a systematic review. *Complement Ther Med*. 2010;18(1):49-56. 10.1016/j.ctim.2009.10.001

68. Latthe PM, Champaneria R, Khan KS. Dysmenorrhoea. *BMJ Clin Evid*. 2011;2011:21.

69. Chung YC, Chen HH, Yeh ML. Acupoint stimulation intervention for people with primary dysmenorrhea: Systematic review and meta-analysis of randomized trials. *Complement Ther Med*. 2012;20(5):353-63. 10.1016/j.ctim.2012.02.008

70. Chen MN, Chien LW, Liu CF. Acupuncture or Acupressure at the Sanyinjiao (SP6) Acupoint for the Treatment of Primary Dysmenorrhea: A Meta-Analysis. *Evid Based Complement Alternat Med*. 2013;2013:493038. 10.1155/2013/493038

71. Jiang HR, Ni S, Li JL, Liu MM, Li J, Cui XJ, et al. Systematic review of randomized clinical trials of acupressure therapy for primary dysmenorrhea. *Evid Based Complement Alternat Med*. 2013;2013:169692. 10.1155/2013/169692

72. Kannan P, Claydon LS. Some physiotherapy treatments may relieve menstrual pain in women with primary dysmenorrhea: a systematic review. *Journal of Physiotherapy (Elsevier)*. 2014;60(1):13-21.

73. Abaraogu UO, Tabansi-Ochuogu CS. As Acupressure Decreases Pain, Acupuncture May Improve Some Aspects of Quality of Life for Women with Primary Dysmenorrhea: A Systematic Review with Meta-Analysis. *J Acupunct Meridian Stud*. 2015;8(5):220-8. 10.1016/j.jams.2015.06.010

74. Song HJ, Seo HJ, Lee H, Son H, Choi SM, Lee S. Effect of self-acupressure for symptom management: a systematic review. *Complement Ther Med*. 2015;23(1):68-78. 10.1016/j.ctim.2014.11.002

75. Abaraogu UO, Igwe SE, Tabansi-Ochiogu CS. Effectiveness of SP6 (Sanyinjiao) acupressure for relief of primary dysmenorrhea symptoms: A systematic review with meta- and sensitivity analyses. *Complement Ther Clin Pract*. 2016;25:92-105. 10.1016/j.ctcp.2016.09.003

76. Smith CA, Armour M, Zhu X, Li X, Lu ZY, Song J. Acupuncture for dysmenorrhoea. *Cochrane Database Syst Rev*. 2016;4(4):CD007854. 10.1002/14651858.CD007854.pub3

77. Armour M, Smith CA, Steel KA, Macmillan F. The effectiveness of self-care and lifestyle interventions in primary dysmenorrhea: a systematic review and meta-analysis. *BMC Complement Altern Med*. 2019;19(1):22. 10.1186/s12906-019-2433-8

78. Ghiasi A, Keramat A, Mollaahmadi L, Hashemzade M, Bagheri L. The effect of acupressure at the Sanyinjiao (SP6) point on relief of primary dysmenorrhea: a systematic review of clinical trials. 2020;19:55-68.

79. Bourdel N, Alves J, Pickering G, Ramilo I, Roman H, Canis M. Systematic review of endometriosis pain assessment: how to choose a scale? *Hum Reprod Update*. 2015;21(1):136-52. 10.1093/humupd/dmu046

80. Hassan Z, Abbas Chishty M, Azam Y, Razzaq T, Shahzadi N, Saeed S. Menstrual Distress among School and College Students. *Ann Med Health Sci Res*. 2021;11:S4:1-4.

81. Moghadam AD, Delpisheh A, Rezaeian M, Khosravi A. Factors Affecting the Labor : A Review Article. *Biomedical & Pharmacology Journal*. 2013;6(2):161-7. 10.13005/bpj/399

82. Mollart LJ, Adam J, Foureur M. Impact of acupressure on onset of labour and labour duration: A systematic review. *Women Birth*. 2015;28(3):199-206. 10.1016/j.wombi.2015.03.007

83. Makvandi S, Mirzaiinajmabadi K, Sadeghi R, Mahdavian M, Karimi L. Meta-analysis of the effect of acupressure on duration of labor and mode of delivery. *Int J Gynaecol Obstet*. 2016;135(1):5-10. 10.1016/j.ijgo.2016.04.017

84. Smith CA, Armour M, Dahlen HG. Acupuncture or acupressure for induction of labour. *Cochrane Database Syst Rev*. 2017;10(10):CD002962. 10.1002/14651858.CD002962.pub4

85. Najafi F, Jaafarpour M, Sayehmiri K, Khajavikhan J. An Evaluation of Acupressure on the Sanyinjiao (SP6) and Hugo (LI4) Points on the Pain Severity and Length of Labor: A Systematic Review and Meta-analysis Study. *Iran J Nurs Midwifery Res*. 2018;23(1):1-7. 10.4103/ijnmr.IJNMR_184_15

86. Smith CA, Collins CT, Levett KM, Armour M, Dahlen HG, Tan AL, et al. Acupuncture or acupressure for pain management during labour. *Cochrane Database Syst Rev*. 2020;2(2):CD009232. 10.1002/14651858.CD009232.pub2

87. Bejar C, Santiago M. Oral. *J Matern Fetal Neonatal Med*. 2016;29 Suppl 1:1-313. 10.1080/14767058.2016.1191212

88. Gameiro L, Barros M, Risso S. The effect of acupressure on pregnant women during labor: A systematic review of literature. *Atencion Primaria*. 2016;48 (Supplement C):21-2.

89. Keramat A, Ganji G, Shiravani MA. Effect of acupressure on labor pain relief: A systematic review of clinical trials. [Persian]. *Iranian Journal of Obstetrics, Gynecology and Infertility*. 2014;17(119):8-17.

90. Mafetoni RR, Shimo AKK. O uso da acupressão para evolução do trabalho de parto e alívio da dor. *Cogitare enferm*. 2013;18(2):365-71.

91. Santiago-Vasco M, Bejar-Poveda C. Meta-analysis of the effects of acupressure during labor. [Spanish]. *Matronas Profesion*. 2017;18(2):60-7.

92. Department of Health. Prolonged pregnancy 2020 [cited 5 November 2021]. Available from: https://www.health.gov.au/resources/pregnancy-care-guidelines/part-j-clinical-assessments-in-late-pregnancy/prolonged-pregnancy.

93. Smith CA, Armour M, Dahlen HG. Acupuncture or acupressure for induction of labour. *Cochrane Database of Systematic Reviews*. 2017(10). 10.1002/14651858.CD002962.pub4

94. Wei CK, Leng CY, Siew Lin SK. The use of the Visual Analogue Scale for the assessment of labour pain: a systematic review. *JBI Libr Syst Rev*. 2010;8(24):972-1015. 10.11124/01938924-201008240-00001

95. Zhang EW, Jones LE, Whitburn LY. Tools for assessing labour pain: a comprehensive review of research literature. *Pain*. 2023;164(12):2642-52. 10.1097/j.pain.0000000000003004

96. Hewitt V, Watts R. The effectiveness of non-invasive complementary therapies in reducing postoperative nausea and vomiting following abdominal laparoscopic surgery in women: a systematic review. *JBI Libr Syst Rev*. 2009;7(19):850-907. 10.11124/01938924-200907190-00001

97. Lee A, Chan SK, Fan LT. Stimulation of the wrist acupuncture point PC6 for preventing postoperative nausea and vomiting. *Cochrane Database of Systematic Reviews*. 2015;2015(11):CD003281. https://dx.doi.org/10.1002/14651858.CD003281.pub4

98. Zhu K, Shen Y, Zhu Y, Li L, You Y. Effect of acupoint therapies on prostatitis: A systematic review protocol. *Medicine (Baltimore)*. 2020;99(6):e18967. 10.1097/MD.0000000000018967

99. **方茜**, **郭霏**, **叶春香**. **穴位按摩减**轻成人腹腔镜术后胃肠道并发症效果的Meta**分析**. *中西医结合护理*. 2019;5(12):43-8. 10.11997/nitcwm.2019012012

100. Zhang GL, Yang SY, Zhu ZL, Mu PX. Meta-analysis on postoperative complications of wristband acupoint pressure therapy. *J Biol Regul Homeost Agents*. 2015;29(1):187-93.

101. Yang J, Jiang Y, Chen Y, Sun M, Chen J, Zheng Q, et al. Acupressure the PC6 point for alleviating postoperative nausea and vomiting: A systematic review protocol. Medicine (Baltimore). 2019;98(33):e16857. 10.1097/MD.0000000000016857