Project

Shiatsu for preventing and treating health conditions: a protocol for an evidence evaluation

Prepared for

National Health and Medical Research Council

NHMRC | Natural Therapies Working Committee
Canberra ACT 2601

CONFIDENTIAL

Research Protocol prepared by
Health Technology Analysts Pty Ltd

March 2021

Contents

Protocol information

Authors

Jorgensen MA[1], Allerdice S[1], Baidya S[1], Shikhule A[1], Miles A[1]

1Health Technology Analysts, Level 1, 370 Norton Street, Lilyfield, New South Wales 2040 Australia

Contact person

Please send all correspondence to Health Technology Analysts Pty Ltd

mailing address: Level 1, 370 Norton Street, Lilyfield, New South Wales 2040 Australia

email: natural.therapies@htanalysts.com.au

Dates

This Research Protocol received approval from the National Health and Medical Research Council (NHMRC) Natural Therapies Working Committee on 11 March 2021.

History

The NHMRC has been engaged by the Department of Health (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 seven natural therapies to be reviewed in the first tranche are naturopathy, Pilates, Rolfing, shiatsu, Tai Chi, Western herbalism and yoga. These therapies are among those excluded from the private health insurance rebate as of 1 April 2019.

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

This Research Protocol has been developed by HTAnalysts, in conjunction with the NHMRC, the NTWC and the NTREAP to provide a framework outlining the methodology that will be used to supplement evidence statements made in the 2015 Review of shiatsu. It is intended that all associated materials will be developed in a robust and transparent manner in accordance with relevant best practice standards (2, 3).

List of tables

List of figures

List of abbreviations

AMSTAR	A MeaSurement Tool to Assess systematic Reviews
BRISA	Regional Base of Health Technology Assessment Reports of the Americas
CI	Confidence interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
COMET	Core Outcome Measures in Effectiveness Trials
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HR	Hazard ratio
MD	Mean difference
MeSH	Medical Subject Headings
NHAA	National Herbalists Association of Australia
NHMRC	National Health and Medical Research Council
NRSI	Nonrandomised study of an intervention
NTREAP	Natural Therapies Review Expert Advisory Panel
NTWC	Natural Therapies Working Committee
OR	Odds ratios
PAHO	Pan American Health Organization
PICO	Population, Intervention, Comparator, Outcome
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	randomised controlled trial
RoB	Risk of bias
RR	Risk ratio
SD	Standard deviation
SMD	Standardised mean difference
SR	Systematic review
TIDIER	Template for Intervention Description and Replication

1 Background

In 2015, a review of shiatsu commissioned by NHMRC found no clear evidence demonstrating its efficacy in treating any clinical condition (1). The 2015 Overview was underpinned by an overview of systematic reviews (SRs) that focused solely on shiatsu and were published in the English language between 2008 and June 2014. Randomised controlled trials (RCTs) that were reported within included SRs and assessed shiatsu delivered to treat any clinical condition were eligible for inclusion; however, no evidence was found (4, 5). SRs that assessed acupressure as a sole intervention were not eligible for inclusion. The 2015 Overview informed the 2015 Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies, which resulted in shiatsu and 15 other natural therapies being excluded from private health insurance rebates[a].

In this 2020 review, the evidence evaluation will build upon the 2015 Overview but will not be limited by publication date and a broader range of study types will be eligible for inclusion (inclusive of nonrandomised studies of interventions [NRSIs]). The updated review will also include studies that assess shiatsu for primary prevention. Evidence from eligible primary studies of shiatsu will be supplemented with evidence identified from an Overview of Reviews (a systematic review of systematic reviews) of acupressure, a core component of shiatsu.

Similar to the 2015 review, eligible comparisons will be shiatsu versus control (but further delineated to shiatsu versus placebo and shiatsu versus no intervention) and shiatsu versus other interventions. Studies not published in the English language will not be translated, and databases in languages other than English will not be searched.

The process for conducting the review is built upon the following framework:

1. source the clinical evidence by performing a systematic literature search of the literature,
2. identify eligible studies published in English and indexed in English language databases,
3. incorporate additional literature identified through non-database sources received from the Department's public call for evidence, NTREAP and NTWC,
4. critically appraise and present the evidence, and
5. determine the certainty in the evidence base for each question, using a structured assessment of the body of evidence in accordance with GRADE methodology (6).

[a] https://www.health.gov.au/resources/publications/private-health-insurance-reforms-changing-coverage-for-some-natural-therapies

1.1 Description of condition and setting

Shiatsu therapy has been practiced in Australia since the early 1970s and is used by people of all ages for a wide of range of reasons (7). A 2016 workforce survey of 121 shiatsu therapists working in Australia found the most common reasons for providing shiatsu therapy were to alleviate or treat a broad range of clinical and preclinical conditions including stress, mental health, pain, musculoskeletal problems, as well as rehabilitation and management of chronic health conditions including cancer care and disability care (7). Other conditions that practitioners report benefit from shiatsu therapy include headaches, migraine, sciatica, respiratory illnesses, fatigue, menstrual problems, circulatory problems, and rheumatic and arthritic complaints (8).

The current review is not limited to any particular condition or setting (see 3.1.2 Types of participants) and therefore a concise description of each condition or problem, and the relevant setting, will be provided after conduct of the full text review.

1.2 Description of intervention

Shiatsu

Shiatsu therapy is a complex, whole-system intervention that is based on the philosophy and theory of traditional Chinese (oriental) medicine (9). Shiatsu was popularised in the West following the legislation in Japan of "Anma-Shiatsu-Massage" in the mid-20th Century. The word 'shiatsu' originates from the Japanese word meaning 'finger pressure', through which the therapist applies pressure to the acupuncture (tsubo) points on the body. Practitioner or self-administered acupressure is believed to assist with alleviating a variety of symptoms associated with health conditions (10).

In Australia, shiatsu therapists work in a range of settings: solo or group shiatsu clinics, multidisciplinary health clinics, at their home or in a mobile practice that allows for home or office visits (7). Shiatsu treatment is traditionally performed on a futon on the floor with the recipient fully clothed or covered by a sheet in a sitting or lying position (11).

Shiatsu therapists working in Australia may have differing qualifications (including some that also receive training in acupuncture and other oriental therapies) and be registered as a shiatsu therapist and/or oriental massage therapist with one or more of the Australian natural therapy associations (7). The Shiatsu Therapy Association of Australia (STAA), a national peak body representing shiatsu therapists, requires practitioners to hold a Diploma in Shiatsu and Oriental Therapies HLT52215 (or equivalent), to undertake a minimum of 20 hours continuing professional education annually and hold a current first aid certificate (12). The same (HLT52215) course may also be recognised by other associations as qualifying for Oriental (TCM) Remedial Massage Therapy. The HLT52215 curriculum covers a range of shiatsu and traditional oriental/Chinese medicine modalities. While the curriculum focuses on the application of acupressure and massage, students are also trained in traditional Chinese medicine diagnosis and the application of moxibustion, cupping, self-acupressure, oriental diet, corrective exercises, lifestyle, relaxation, breathing techniques and meditation (7).

Shiatsu draws on traditional Chinese philosophies and theories such as Yin and Yang, the energy meridians, the five elements and the concept of Ki. The therapist aims to identify and correct imbalances in the flow of Ki, or energy in a recipients meridians to promote self-healing (11, 13). A shiatsu session typically begins with a clinical history followed by an examination to formulate a traditional Chinese medicine diagnosis that is used to individualise the therapy (11, 13, 14).

Examination methods include palpation (ampuku) of the hara (abdomen) or back, pulse, tongue and face diagnosis to assess the relative constitutional and energetic qualities of the internal organs and their related meridians (11).

Shiatsu massage is a derivative of traditional Japanese (Anma or Amna) massage that has its origins in traditional Chinese (Tuina) massage. There are various styles of shiatsu practiced in Australia and include: barefoot (macrobiotic) shiatsu, healing shiatsu, Jin Shin Do, Namikoshi (or Nippon shiatsu), ohashiatsu, quantum shiatsu, Tao shiatsu, tsubo therapy, Zen shiatsu, and watsu (water shiatsu). Like Anma and Tuina massage, shiatsu massage employs a wide range of traditional techniques. Some are vigorous (e.g. kneading, rubbing, tapping and shaking) while other techniques involve slow and firm or quick and gentle movements (e.g. pressure, stroking, stretches, rotations of the limbs and joints, simple structural alignments and muscle release techniques) (9, 11). All styles, however,

have commonality in their underpinning traditional Chinese medicine and the use of body weight and physical pressure in one way or another when delivering a shiatsu massage (15).

Under standard terminologies published by the WHO (16), 'Traditional Chinese Tuina' is the term used when referring to massage therapy that has its foundations in traditional Chinese medicine, including shiatsu. The term shiatsu, however, is also listed as an alternate word for acupressure, demonstrating that the two terms may be used interchangeably within the literature. Terminology is also specified for a number of components of practice used by shiatsu practitioners in Australia (see Table 1).

Table 1 Description of common components of practice used in shiatsu therapy

Component of practice taught and used by certified Australian shiatsu therapists (HLT52215)[a]	Corresponding WHO traditional medicine terminology (TRM)[b]	WHO TRM Codes
Shiatsu massage	**Traditional Chinese Tuina:** the branch of traditional Chinese medicine concerned with the principles and clinical use of Tuina (massage) therapy	0.0.18
Shiatsu massage techniques	**Massage:** rubbing, kneading, or percussion of the soft tissues and joints of the body with the hands, usually performed by one person on another, esp. to relieve tension or pain	4.2.326
	Manipulation: the term used when describing a wide range of manual therapy techniques (see manual for details)	4.2.327 – 4.2.344
	Tapping technique: a manipulation performed by tapping with the tips of fingers held together	4.2.346
Acupressure	**Finger Pressure:** a manipulation performed by pressing acupuncture points with the finger or thumb instead of needling, also known as shiatsu	4.2.347
Moxibustion	**Moxibustion:** a therapeutic procedure involving ignited material (usually moxa) to apply heat to certain points or areas of the body surface	5.2.0 – 5.2.39
Cupping	**Cupping:** suction by using a vacuumised cup or jar	5.3.0 – 5.3.11

Source:

a. https://training.gov.au/Training/Details/HLT52215 (see Core units HLTSHU001, HLTSHU002, HLTSHU003, HLTSHU004, HLTSHU005, HLTSHU006, HLTSHU007, HLTSHU008, HLTSHU009)

b. World Health Organization international standard terminologies on traditional medicine in the Western Pacific Region (16)

Acupressure

Acupressure is the application of pressure on the traditional Chinese medicine acupuncture (tsubo) points (10, 17). Acupressure is applied to specific points by the use of finger, hand, elbow, foot, and/or acupressure band or bead (10). Like shiatsu, acupressure has its roots in traditional Chinese medicine. However, acupressure is only one component of shiatsu therapy. Some authors note that while similar, there are differences in both the technique and philosophy of shiatsu and acupressure (10, 14, 17). While others, including the WHO International Standard Terminologies for Traditional Medicine in the Western Pacific Region (16), use the terms interchangeably (7, 17, 18).

For the purpose of this review, individualised acupressure refers to the selection of one or more acupressure (tsubo) points tailored to the individual recipient that is informed by a traditional Chinese medicine diagnosis. It may be delivered in the context of a shiatsu treatment (or alongside

other traditional Chinese medicine treatment). Most individualised acupressure is administered by the therapist, however, there may be instances when the therapist provides instructions for self-administration or administration by a lay caregiver.

Standardised (non-individualised) acupressure refers to one or more predetermined acupressure points for a specific condition, symptom or symptom cluster. While informed by the principles of traditional Chinese medicine, a traditional diagnosis of the recipient is not required for administration. The acupressure intervention may be administered by a qualified therapist, lay caregiver or self. Most self-administered acupressure is non-individualised, examples include the use of an acupressure band or large bead as a replacement for finger pressure. However, there may be some instances when a therapist uses this type of acupressure or provides instructions to their patient or lay caregiver for self-administration.

1.3 How the intervention might work

Shiatsu

The exact mechanisms for how shiatsu might work for different conditions has not been elucidated. However, some of the massage techniques included in shiatsu are similar to massage more generally. Consequently, shiatsu massage techniques may affect the body through similar mechanisms. The physical and psychological effects of massage more generally are thought to be explained (in part) by relaxation of the nervous system and musculature and through the release of neurotransmitters and hormones (19-21). Like all human interaction, touch is believed to induce an immediate calming effect, mediated by production of the hormone oxytocin (22, 23). Physiologically, massage may increase oxygenation to the muscles through inducing local biochemical changes that help blood and lymph flow, as well as affecting mood and pain perception through a subsequent effect on neural activity (20, 24-26). One study suggests that by supporting individuals to take control of their self-care, shiatsu practitioners encourage patients to make lifestyle changes (e.g. relaxing more, working less) which could impact their health (27).

A limited number of exploratory studies of shiatsu applied to various anatomic regions on healthy adults have observed mixed results on autonomic nervous system functions including changes in heart rate, blood pressure and pupil diameter (20). In shiatsu, the role of connective tissue in delivering energy and information to the whole body is thought to be integral to correcting imbalances in the flow or energy of a recipients Ki (8). Here, the mechanical effect of stretching and applied pressure is believed to encourage hydration and flexibility of connective tissue that has become congested or brittle through inadequate circulation or after injury. Due to continuity and electromagnetic signalling properties between cells, the connective tissue is suggested to distribute and enhance communication throughout the body to promote self-healing (8). Shiatsu practitioners work directly with the energy flow along the meridian because bioelectric flow is believed to happen more readily along meridian lines due to differing electrical resistance (8).

Acupressure

The mechanisms underpinning any effects of acupressure are thought to be similar to acupuncture. The practice of acupuncture may produce benefits for conditions closely associated with the nervous system (28-30). Stimulation (or pressure) applied to points connected to or located near neural structures is thought to transmit information along pathways in the nervous system, inducing various changes including pain perception, and hormonal and neurochemical changes (28-30).

1.4 Why it is important to do this review

In Australia, complementary therapies, including shiatsu, are most often used in conjunction with conventional medicine and other strategies for maintaining good health and wellness. For this reason, it is important to synthesise the evidence for the effectiveness of shiatsu, to enable consumers, health care providers and policy makers to make informed decisions about care.

The 2015 Overview identified four systematic reviews that included two RCTs; however, the effect of shiatsu could not be independently evaluated from either RCT, as shiatsu was combined with other interventions in both studies. There was no independent evidence to suggest that shiatsu may have some beneficial health benefits compared with control or active comparators. Therefore, the overall efficacy, safety, quality or cost-effectiveness of shiatsu was not able to be assessed (4).

The rationale for conducting this review is to update and enhance the evidence used to inform the 2015 Overview of shiatsu (4). That is, to identify studies published since, or not included in, the 2015 review, and address the evidence gaps noted. This is to ensure recommendations relating to the use of shiatsu remain relevant and up to date.

2 Objectives

To conduct a systematic review of RCTs and NRSIs to evaluate the effectiveness of shiatsu in individuals with a described injury, disease, medical condition, or preclinical condition. This will be supplemented with the available evidence for acupressure – a core component of shiatsu - as reported in SRs (of RCTs and pseudo-RCTs).

The intent is to evaluate the evidence representative of the populations and conditions commonly seen by shiatsu practitioners in Australia, the intervention(s) commonly used by practitioners, and outcomes that align with the reasons why patients use shiatsu and/or practitioners prescribe shiatsu.

3 Methods

Methods reported in this protocol are based on that described in the *Cochrane Handbook for Systematic Reviews of Interventions* (31) and relevant sections in the Joanna Briggs Institute Reviewer's manual (32). Covidence (www.covidence.org), a web-based platform for producing systematic reviews, will be used for screening citations and recording decisions made. Covidence is compatible with EndNote and Microsoft Excel, which will be used for managing citations and data extraction, respectively. Where appropriate, RevMan (33) will be used for the main analyses and GRADEpro GDT software (www.gradepro.org) will be used to record decisions and derive an overall assessment of the certainty of evidence for each outcome guided by GRADE methodology (6). The final approved review protocol is to be registered on the international prospective register of systematic reviews (PROSPERO).

To identify the evidence base for the clinical question a systematic search of published medical literature will be conducted. All potentially relevant studies will be identified after applying prespecified inclusion and exclusion criteria as outlined below.

3.1 Criteria for considering studies for this review

3.1.1 Types of studies

Study design

Eligible studies are those designed to examine the effectiveness of shiatsu, or its core component – acupressure – compared to control (placebo or no intervention), or other intervention. The evidence for acupressure and shiatsu will be examined separately. For shiatsu, eligible studies are RCTs and NRSIs. For acupressure, eligible studies are SRs (of RCTs and pseudo-RCTs).

Shiatsu

The primary study of interest is an RCT. Cluster-randomised trials and crossover trials are also eligible for inclusion, and will be analysed using methods appropriate to the design (see Section 3.3.9) (34). If the method of randomisation is not specifically stated, or not considered strictly random, then the study will be judged to be pseudorandomised. Pseudo-RCTs will be evaluated alongside RCTs, with methods of randomisation examined in the risk of bias assessment and any concerns about risk of bias addressed in the synthesis.

Certain NRSIs with design features as outlined in Table 2 are also eligible for inclusion. For NRSIs to be eligible for inclusion, the minimum design features include:

- allocation to, or practice of, the intervention occurs by choice (by the participant or other)
- the effect of the intervention in individuals (or clusters of individuals or groups) is compared with a contemporaneous control group

NRSIs in which the effect of the intervention is compared to a historical (or non-parallel or non-concurrent) control group are not eligible for inclusion due to concerns regarding risk of bias (e.g. due to residual confounding or unmeasurable changes in clinical practice over time). Single arm studies with either post-test or pre-test/post-test outcomes, cross-sectional studies, case series and case reports are also not eligible for inclusion, as it is too problematic to assess the effect of the intervention in such studies with any confidence (35, 36) .

The inclusion of NRSIs is to ensure the evidence review adequately covers the breadth of health conditions and outcomes to inform health policy, particularly in populations or settings where the intervention is either not likely, or not able, to be assessed using a randomised design, or where evidence from RCTs is incomplete for certain populations, settings or outcomes that may be more feasibly evaluated using NRSI (37). This is likely to occur when the length of follow up for the outcome is not feasible for an RCT, or the event rate of the outcome is so small that it requires a population-wide study for a measurable effect to be observed. In rare instances, it may be because of a strong preference for the intervention by prospective participants prevents the conduct of a suitable RCT (38), or the RCT evidence for a particular health condition and outcome is indirect and the question is better answered by available NRSI evidence (39).

Eligible NRSIs that are assessed to be at critical risk of bias for one or more domain (see Section 3.3.6) will not be included in the evidence synthesis because results from these studies are likely to lead to misinformed judgements about the effect estimate.

Table 2 Eligible design features of nonrandomised studies of interventions

	Definition / design features
Design features of NRSIs included in the review	An experimental study in which people are allocated to the intervention/treatment being studied or a control/placebo group and the outcomes compared. The method of allocation is by choice, availability, or chance.
	A study in which outcomes from a defined group of people (the cohort) are followed over time, to examine associations between exposure and non-exposure to an intervention or factor under study. Outcome are recorded as they occur. A 'prospective' cohort study recruits participants before any intervention and follows them into the future.
	A study in which outcomes from a defined group of people (the cohort) are identified to examine associations between exposure and non-exposure to an intervention or factor under study. A 'retrospective' cohort study identifies subjects from past records describing the interventions received and follows them from the time of those records.
	A study that uses observations at multiple time points before and after an intervention (the 'interruption') is introduced to a group of people, and then compared to the outcomes at the same time points for a group of people that do not receive the intervention. The design attempts to detect whether the intervention has had an effect significantly greater than any underlying trend over time.
	A study in which observations are made before and after the implementation of an intervention, both in a group that receives the intervention and in a control group that does not and compared at the same timepoint.
	A study that compares people with a specific outcome of interest ('cases') with people from the same source population but without that outcome ('controls'), to examine the association between the outcome and prior exposure (e.g. having an intervention). This design is particularly useful when the outcome is rare.

Source: Adapted from NHMRC (36, 40); Chapter 24 Including non-randomised studies on intervention effects (37); Cochrane Childhood Cancer (41);

Acupressure

The primary study of interest is a SR of RCTs (and pseudo-RCTs), with or without a meta-analysis. This is because conducting a SR of primary studies of acupressure (i.e. RCTs or NRSIs) is not feasible given the timeframe and resources. As acupressure is a central technique used in shiatsu, evidence from SRs of acupressure will be used to augment the evidence from primary studies of shiatsu.

If the method of randomisation of a primary study included within a SR is not specifically stated, or not considered strictly random, then the study will be considered to be pseudorandomised. Where a SR includes pseudo-RCTs, these will be considered as eligible along with data from RCTs. Reviews that do not report study eligibility criteria or conduct a comprehensive search of the literature (i.e. searching more than one database) will not be included. These reviews do not meet the minimum criteria to be considered 'systematic' and may not accurately summarise the body of evidence.

Eligible reviews that include a single RCT will be included, as will SRs that include both RCTs and NRSIs; however, only evidence from the RCTs (or pseudo-RCTs) will be considered. Information on how meta-analyses from SRs will be handled where they include ineligible studies (e.g. NRSIs that are not pseudorandomised) or where they are missing one or more eligible studies is provided in Section 3.3.11.

Additional study designs will not be considered. This includes individual RCTs or pseudo-RCTs not part of a SR, nonrandomised comparative studies (i.e. nonrandomised experimental trials, cohort studies, case-control studies, interrupted times series), cross-sectional studies and case series with

either post-test of pre-test/post-test outcomes.. Where a SR is not identified for an eligible population or intervention, this will be noted as an evidence gap.

Overviews will not be eligible for inclusion and the search strategy (see Appendix A) is not specifically designed to identify them. However, Overviews identified in literature search or those submitted through the Department's public call for evidence (see Section 3.2.2) will be checked to identify SRs that are eligible.

Publication date

There are no limitations on publication date, however, studies or systematic reviews published after the literature search date will not be eligible for inclusion. Studies that are published (or submitted to the Department) after the literature search date will be listed within the '*Studies awaiting classification*' table of the evaluation report. These studies will not be subject to a formal evidence evaluation, however, a brief statement about the study and its potential impact on the overall conclusions of the evidence review will be included under the relevant sections of the review (e.g. '*Overall completeness and applicability of evidence*').

Studies published in languages other than English

The literature search, as well as the Department's call for evidence, will not be limited by language of publication. Databases in languages other than English will not be searched, however, studies in languages other than English may be identified via the English language databases. For pragmatic reasons, potentially eligible studies will not undergo full text translation or data extraction, but will be documented via a process outlined in Section 3.3.1 *'Studies published in languages other than English'*.

3.1.2 Types of participants

People of any age with any injury, disease, medical condition or preclinical condition are eligible for inclusion. This includes disease prevention in at-risk healthy populations, which is broadly defined as those who are at increased risk of becoming ill or injured based on social, biomedical or behavioural risk factors (42). For the purposes of this review, social risk factors include income, education, employment and social support; biomedical factors include a person's age, genetic make-up, and health status (such as obesity, high blood pressure, high cholesterol, vitamin deficiency); and behavioural factors include a person's lifestyle choices (e.g. alcohol consumption, diet, exercise, tobacco and other drug use, etc.).

Healthy participants seeking health improvement, such as general wellbeing, fitness, aesthetic improvements, resilience and cognitive or emotional intelligence are not eligible for inclusion; however, a study with eligible and ineligible populations will be included if separate data are available for the eligible population/s.

3.1.3 Types of interventions

Intervention: shiatsu

All styles and forms of massage described as shiatsu or a type of shiatsu are eligible for inclusion. Other components of practice used by shiatsu therapists in Australia (such as moxibustion, cupping, self-acupressure, oriental diet, corrective exercises, lifestyle, relaxation, breathing techniques and meditation) are only included if they are delivered in the context of whole-system/multi-component shiatsu therapy. Shiatsu may be individualised according to a traditional Chinese medicine diagnosis or a predetermined standardised intervention for a specific condition/problem.

Terminology specific to the Japanese language is not used in the standardised nomenclature (16). Therefore, studies using other traditional oriental terms to describe the intervention are also included in the review. This includes, but is not limited to, shiatsu massage, Anma (or Amna) massage, An mo massage, Tsubo (acupoint) therapy. Studies described as 'Traditional Japanese massage' or 'Japanese massage' will be included where the description of the intervention is consistent with the definition of shiatsu.

There are no limits on intensity, duration of intervention or mode of delivery and studies will be included irrespective of whether the intervention is delivered by a certified practitioner or not.

To allow for potential subgroup analysis (and to inform decision-making), studies will be stratified based on the style or form of shiatsu massage delivered (see Section 3.3.15).

Restrictions: Therapies that are described as Tuina (or tui na) or other styles of massage or manual therapies (e.g. Korean massage, traditional Chinese massage, remedial massage, Swedish massage etc.) are excluded. Interventions that include high-velocity joint manipulation or oils are excluded, as these components are not consistent with shiatsu practice in Australia. Studies only evaluating component interventions, such as acupressure[b], moxibustion, cupping and meridian exercises, will be excluded.

Intervention: acupressure

All forms of acupressure, applied to traditional Chinese medicine acupuncture (tsubo) points, are eligible for inclusion. The acupressure can be either individualised (i.e. delivered according to a traditional Chinese medicine diagnosis or a predetermined condition/problem) or non-individualised (i.e., delivered to predetermined acupressure points for a specific condition/symptom) and can be applied by a therapist to specific points using a finger, hand, elbow, or foot. The acupressure can also be self-administered or delivered by a lay caregiver and/or via an acupressure band or large bead as a replacement for finger pressure. Studies in which the acupoint is stimulated by needles or other techniques (i.e. by pressure/laser or transcutaneous electrical nerve stimulation) are not eligible for inclusion.

There are no limits on intensity, or duration of intervention and studies will be included irrespective of whether the intervention is delivered by a certified acupressure or shiatsu practitioner or not.

Studies that include acupressure delivered alongside shiatsu massage and/or as a component of what is described as whole-system shiatsu therapy will be considered in the shiatsu component of the review (see Section 3.3.11).

SRs that consider a broader question than intended for this review (e.g. assesses the effectiveness of acupressure among other interventions) will be included if the SR specifically assesses the effectiveness of acupressure independent of the other included interventions. If only a subset of studies contained within the SR meet the eligibility criteria for this review, then only those eligible primary studies as reported in the SR will be considered (see Section 3.3.11).

Restrictions: Studies that only evaluate auricular acupressure will be excluded, due to indirectness with shiatsu.

[b] Primary studies evaluating acupressure are excluded from the shiatsu component of the review (SR of RCTs and NRSIs). However, SRs of acupressure will be considered within the Overview of Reviews (SR of SRs) component of the review.

Comparators

There are no restrictions on the type of eligible comparators, noting that the analysis will stratify the evidence into three comparisons: (i) placebo; (ii) no intervention, wait list or usual care (unless active); and (iii) other interventions (inclusive of usual care if considered active). The decision to separate placebo and no intervention has been made to account for any potential placebo effect which may occur. For instance, while sham interventions are designed to be a placebo, some have demonstrable clinical effects (43).

Where usual care is poorly described or where usual care is described as an adjunct to shiatsu it will be considered an inactive intervention. Other comparators could include (but will not be limited to) pharmacologic treatments, manual therapies, exercise programs or other forms of physical activity designed to improve health.

Co-interventions such as diet, education programs, lifestyle modification, or medication may be administered simultaneously to the intervention and comparison group. Studies with co-interventions that are not provided within the context of shiatsu therapy will be included if all arms of a study receive the same co-interventions (i.e. the effectiveness of shiatsu is not confounded).

Restrictions: Studies comparing different styles, forms, or components of shiatsu with one another (including studies comparing shiatsu with acupressure) will be excluded.

3.1.4 Types of outcome measures

Outcome role

Outcomes will not be used as a criterion for including studies (or SRs).

Outcome domains of interest

Outcomes are intended to align with the reasons why patients use the therapy and/or practitioners prescribe the therapy. This includes recovery, rehabilitation, and changes in disease outcomes and symptoms (e.g. pain, joint range of motion, strength, balance, and accepted surrogate outcomes such as HbA1C for diabetes, body mass index for weight gain or loss, lung function tests), health related psychological/behavioural outcomes, health related quality of life, self-reported benefits, symptoms and functional ability, medication use or compliance with conventional medicine treatment; and injury or disease specific prevention outcomes (e.g. falls prevention, smoking cessation).

Restrictions: Consistent with the terms of reference of NTREAP, personal health care preferences, patient-reported experience measures (PREMS) (e.g. satisfaction with care), safety, quality and economic outcomes are out of scope.

Outcome measures and time points of interest

Any effectiveness outcome anticipated to demonstrate a treatment achieves its intended purpose is eligible for inclusion (35, 36). There are no limitations on time points (e.g. short and long term outcomes) or outcome measure (e.g. objective and subjective measures such as clinical and laboratory assessments and patient-reported outcome measures [PROMS], preferably measured using validated tools, are eligible).

As there are a broad range of populations eligible for inclusion in the review, it is not possible to prespecify outcomes. All prespecified outcomes measured in each eligible SR, RCT or NRSI will be listed in the '*Characteristics of included studies*' tables; however, results will only be extracted for

those outcomes identified as critical or important to the review. For each identified population, results for a maximum of seven critical (or important) outcomes will be reported in GRADE summary of findings (SoF) tables and corresponding evidence statements (see Section 3.3.17).

Outcome selection will occur after identification of eligible studies using a prespecified approach. To avoid introducing bias, outcomes will be prioritised by the NTWC, who will be provided with a list of conditions, outcome domains and outcome measurements (including measurement tools and time points) to prioritise. This list will be derived from the outcomes reported in studies identified for inclusion in the review, and, where available, core outcome set/s for a particular condition (identified by searching COMET).

Throughout the outcome prioritisation exercise, the NTWC will remain blinded about the characteristics or results of included studies (or SRs) to prevent knowledge of study results or other characteristics (such as study design) from influencing decision-making. In determining the critical and important outcomes, the NTWC will be guided by GRADE (6), and focus on the relevance and validity of outcome measures. Where appropriate, outcome domains reported using different measurement tools will be grouped and reported accordingly (see Section 3.3.8).

Outcomes reported at different timepoints will be grouped and considered as follows: short term, intermediate term, long term, or not specified. Determining whether something is considered short, intermediate or long term for a population will be guided by the published evidence, the NTWC and COMET. To avoid unit-of-analysis issues associated with repeated observations (see Section 3.3.9), data from a single time point will be selected for each outcome, as determined by the NTWC during outcome prioritisation. Where multiple timepoints are considered critical or important to decision-making (e.g. short- and long-term remission in symptoms), separate outcomes will be specified for each timepoint.

3.2 Search methods for identification of studies

3.2.1 Electronic searches

The literature search strategy (see Appendix A) was developed in Ovid (for Embase, MEDLINE, and Emcare) based on the key element of research question (i.e. the intervention). The search is not limited by population or outcome, but rather by study type; with methodological filters for identifying SRs, RCTs, and NRSIs and exclusions for publication types developed and published previously (44).

In developing the search strategy, we appraised and adapted the relevant search strategies provided in the 2015 Review; with recent systematic reviews identified in the scoping report and studies suggested by the NTWC also reviewed to identify other potentially relevant concepts. Terms or concepts proven not suitable were removed and other terms added.

No date, language or geographic limitations will be applied when conducting the search. Non-English databases will not be searched.

The search strategy will be adapted to suit the required syntax for the electronic bibliographic databases outlined in Table 3.

Table 3 Electronic bibliographic databases to be searched

Database	Shiatsu	Acupressure
Embase (via OvidSP)	✓	✓
MEDLINE (via OvidSP)	✓	✓
Emcare (via OvidSP) – coverage of all nursing specialty areas	✓	✓
Cochrane Database of Systematic Reviews (via Cochranelibrary.com)	X	✓
Cochrane Central Register of Controlled Trials (via Cochranelibrary.com)	✓	X
PsycINFO (via OvidSP) – coverage of behavioural science and mental health	✓	✓
AMED (via *OvidSP*) – coverage of Allied and Complementary Medicine	✓	✓
CINAHL (via *EBSCOHost*) – Cumulative Index to Nursing and Allied Health Literature	✓	✓
SPORTDiscus (via *EBSCOHost*) – coverage of exercise physiology, medicine, biomechanics, coaching, counselling, psychology, and sports medicine,	✓	✓
PubMed (limited to in-process citations and citations not indexed in MEDLINE) – to retrieve citations not yet indexed in OVID	✓	✓
Systematic Review Data Repository (via the Agency for Healthcare Research and Quality)	X	✓
Pan American Health Organization (PAHO) Virtual Health Library (VHL) – including Lilacs (Health information from Latin America and the Caribbean countries), PAHO IRIS (institutional repository for information sharing), and BRISA (Regional Base of Health Technology Assessment Reports of the Americas)	✓	✓

3.2.2 Other sources

Reference lists of key relevant articles will be checked to identify any additional studies not identified through searches of the primary databases. The public will also be invited by the Department to submit references for published research evidence (not examined in the 2015 Review) through a public call for evidence.

Grey literature is excluded, the exception being evidence reviews commissioned by Australian government bodies and other national or international bodies that are recommended by NTREAP or committee members, noting that systematic reviews of primary studies are only eligible for inclusion where the intervention is specified as ‘acupressure’ (see Section 3.1.3). For ‘shiatsu’, any recommended evidence reviews will be used as a source for identifying primary studies not identified through electronic searches.

3.3 Data collection and analysis

In the first instance the evidence reviews will aim to assess the full breadth of eligible studies. However, where the scope of the review becomes unmanageable (as determined post screening i.e. an unmanageable number of eligible citations/ populations), the scope will be narrowed to focus analysis and synthesis of evidence on populations and conditions relevant to Shiatsu practice in Australia, with a focus on evidence evaluating shiatsu first and foremost, augmented with evidence evaluating acupressure.

NHMRC's NTWC will advise on the populations and conditions of interest after considering a blinded list of all eligible populations and conditions post-screening. Populations and conditions will be selected based on objective data about practice in the Australian context (e.g. practitioner or patient surveys that report reasons for use in Australia), where possible. Evidence for populations and conditions not prioritised for synthesis, will be listed in an evidence inventory to ensure that all eligible evidence is catalogued.

Included studies will be critically appraised, appropriate data extracted into data extraction tables, and the results analysed and summarised into appropriate categories according to identified populations, conditions and comparators. Summary of Findings tables will be developed for up to seven critical and important outcomes, guided by the GRADE framework.

3.3.1 Inclusion decisions

Studies identified in the literature searches

Title/abstract screening

Citations (title/abstracts) retrieved by the literature searches will be imported into EndNote and duplicates removed. Citations will then be imported to Covidence (www.covidence.org), an online tool that streamlines the screening and data extraction stages of a systematic review. Screening will be staged according to study design, starting with screening of SRs (for acupressure), then RCTs (for shiatsu), followed by NRSIs (for shiatsu).

Each citation (titles and abstract) will be screened by one evidence reviewer who will discard ineligible studies (marked as irrelevant and tagged with a reason for exclusion) and retain those with relevant data or information (marked as relevant or maybe). Where there is uncertainty regarding eligibility, a decision will be made through discussion with the lead reviewer, who will either decide to mark the citation as irrelevant or take it through to full text. Citations that are in a language other than English will be tagged and managed as described below under '*Studies published in languages other than English*'.

Full text screening

Full text articles identified for possible inclusion in the evidence synthesis will be retrieved and assessed for inclusion by one reviewer. A prespecified, hierarchical approach, as outlined in Appendix B, will be used to annotate reasons for exclusion, with the results of the study selection process illustrated in a PRISMA diagram. Ineligible studies will be marked with a reason for exclusion and listed in a table in the technical report under '*Characteristics of excluded studies*'. Where there is uncertainty regarding eligibility, a decision will be made through discussion with the lead reviewer. The lead reviewer will also reinspect a random 20% sample of articles marked as excluded to ensure adherence to the *a priori* exclusion criteria and any differences will be resolved by discussion. If additional expertise or advice regarding the application of the PICO criteria is required, further

follow up with the NTWC will occur (noting that NTWC will be presented with excerpts from the publication relevant to the query while remaining blinded to other identifying details such as the study citation, design, size, risk of bias, and results).

If a study (or SR) does not contain the required PICO information for a decision to be made regarding eligibility, the information will be sought from the study's authors through an open-ended request. Trial registration numbers, author names, and study titles, locations and dates will be used to identify multiple reports arising from the same study (or SR). Eligible studies that are not available in English will be noted and managed as described below under '*Studies published in languages other than English'*

Overviews identified in the literature search will not be eligible for inclusion in the review but will be checked for eligible SRs.

Evidence provided through the Department's public call for evidence

Potentially relevant SRs and primary studies identified by the NTWC, NTREAP, and other key stakeholders will be considered for inclusion if they satisfy the eligibility criteria described in Section 3.1 above. Overviews provided through the Department's public call for evidence will be checked for eligible SRs.

All of the submitted literature will be collated, tabulated, and cross-referenced with the evidence identified in the literature search described in Section 3.3.1. In-scope studies (or SRs) not identified in the literature search will be incorporated into the evidence evaluation. A rationale for exclusion (as noted in Appendix B) will be provided for all studies considered out of scope (documented in a table within the technical report).

Studies published in languages other than English

Studies (or SRs) published in languages other than English will undergo title and abstract translation using Google translate (or an equivalent tool). If online translation does not facilitate understanding of the title and abstract, then these studies will be listed in a table as *'Studies unable to be translated or interpreted at the title/abstract stage'*. Translated titles and abstracts will be screened to remove irrelevant citations, with articles excluded at title and abstract screen reported in the *'Results of the search'*.

Translated titles and abstracts will be reviewed and evaluated against the '*Criteria for considering studies for inclusion in the review*'. Full text translation will not occur to determine eligibility. Studies assessed as potentially eligible for inclusion in the review will be recorded in a *'Studies Awaiting Classification'* table. This information will also be reflected in the PRISMA flow diagram.

The potential risk of language bias and its implications for the evidence evaluation will be discussed in relevant sections of the Evaluation Report (such as '*Overall completeness and applicability of evidence*' and '*Agreements and disagreements with other studies or reviews*'). When assessing the extent to which language bias might influence the conclusions, we will also consider whether English-language reviews included in the acupressure component of the review searched for, and included, studies published in languages other than English.

3.3.2 Data collection process

The characteristics of all included primary studies or SRs will be extracted using a standard pre-tested data extraction and coding form (see Appendix D). Outcome data will be extracted after

agreement has been reached regarding the critical and important outcomes to be appraised (see Section 3.1.4). Pre-testing will involve all reviewers, who will data extract the required information from the same two SRs (acupressure) and two primary studies (shiatsu) specifically selected to cover the breadth of the PICO and anticipated study designs identified for inclusion in the review. The lead reviewer will inspect the forms to ensure the relevant data are extracted consistently between reviewers and as planned, with any necessary revisions made to ensure consistency.

All data extraction forms will be checked for completeness and accuracy by the lead reviewer. Where there is uncertainty or disagreement regarding included data, a decision will be made through discussion.

Full data extraction (and critical appraisal) of primary studies included within a SR (for acupressure) will not occur. If a return to primary studies is required to check or confirm information, we will note any discrepancies or adjustments made.

3.3.3 Requests for data

Eligible SR protocols, ongoing trials, studies not published in English, and studies published as conference abstracts will be identified for inclusion. Study authors will be contacted through an open-ended request for data or further information. If no further data are available, the study will be noted as '*Ongoing*' or within the '*Studies Awaiting Classification'* table and will not be included in the evidence appraisal.

No attempts will be made to obtain or clarify data from authors of published peer-reviewed primary studies (shiatsu) or SRs (acupressure).

3.3.4 Data items

Shiatsu

The following characteristics of included studies will be extracted: study design, year conducted, setting and location, participant characteristics (including demographics, comorbidities, etc.), intervention and comparator characteristics (including number of treatment sessions, program duration, co-interventions), outcomes (including measurement method, timing, or severity), method of analysis, and funding sources.

Acupressure

The following characteristics of included SRs will be extracted: review objectives, study design (e.g. qualitative review, meta-analysis), year conducted, databases searched, date (and range) of documented search, SR eligibility criteria for participant characteristics (including demographics, comorbidities, etc.), SR eligibility criteria for intervention and comparator characteristics (including number of treatment sessions, program duration, co-interventions), outcomes reported in the SR (including measurement method, timing or severity), the method of synthesis/analysis employed, characteristics of included primary studies (number, study design features), risk of bias tool used to appraise included primary studies and their rating (noting review authors comments or concerns), funding sources and the overall conclusion of the SR. The data extraction forms will also note whether the SR searched for and included publications in languages other than English.

Included primary studies will also be listed by author, date of publication and eligibility for inclusion in this review. When evaluating the effectiveness of acupressure across the same PICO (See Section 3.3.11and 3.3.12), overlaps and omissions between SRs will be noted in a matrix (see Figure 1).

Where data from a selected SR is augmented with data obtained from another SR this will be documented using footnotes.

3.3.5 Missing data

No imputation for missing data will be conducted. Studies with missing data will be included alongside other studies for that condition; either in the narrative (non-quantitative) synthesis of results or on forest plots showing the sample size. Implications of the missing data will be considered when interpreting the evidence and will be discussed under *'Overall completeness and applicability of evidence'*. Investigations into missing data within a study (e.g. a review of the clinical trial protocol) will be noted when assessing the risk of bias for that study (see Section 3.3.6).

3.3.6 Tools to assess risk of bias in individual studies

The risk of bias of included studies will be assessed using the most appropriate risk of bias assessment tool according to the type of study as follows:

- SRs (acupressure): AMSTAR-2 quality assessment checklist (45)
- RCTs (shiatsu): Revised Cochrane risk of bias tool v2.0 (46, 47)
- NRSIs (shiatsu): ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions (31, 48)

Shiatsu

Randomised controlled trials

The risk of bias of RCTs will be assessed using the revised Cochrane Risk of Bias tool (46, 47). This tool is made up five domains assessing bias arising from the randomisation process; bias due to deviations from intended interventions; bias due to missing outcome data; bias in measurement of the outcome; and bias in selection of the reported result. Each domain will be assessed for bias, which will be recorded as 'high', 'low', or 'some concerns'.

An overall risk of bias for each outcome in the RCT will be judged based on the following criteria:

- *overall low risk of bias* – low risk of bias for all key domains
- *some concerns* – at least one domain has some concerns raised, but none are found to be at high risk of bias
- *overall high risk of bias* – high risk of bias for one or more key domains

Nonrandomised studies interventional studies (NRSIs)

Critical appraisal of NRSIs will be guided by the methods described by Cochrane using the ROBINS-I tool (31, 48). Potential confounders and cointerventions for each population identified for inclusion will be identified and agreed through discussion with the NTWC prior to assessment of the risk of bias. ROBINS-I evaluates the risk of bias observed in the following domains: confounding, selection of participants, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selective reporting. Each domain will be judged, which will be recorded as 'low', 'moderate', 'serious', 'critical', or 'no information provided'.

The overall risk of bias judgement for a specific outcome will use the following guide:

- *overall low risk of bias* – the study is comparable to a well-performed RCT and is judged to be a low risk of bias for ALL domains.

- *overall moderate risk of bias* – the study appears to provide sound evidence for a nonrandomised study but cannot be considered comparable to a well-performed randomised trial. The study is judged to be a low or moderate risk of bias for ALL domains.
- *overall serious risk of bias* – the study has some important problems and is judged to be at serious risk of bias in at least ONE domain, but not a critical risk of bias in any domain.
- *overall critical risk of bias* – the study is too problematic with regards to this domain to provide any useful evidence on the effectiveness of the intervention. The study is judged to be at critical risk of bias in at least ONE domain.
- *no information* – there is no information on which to base a judgement about overall risk of bias. There is no clear indication that the study is at serious or critical risk of bias AND there is a lack of information in one or more key domains of bias.

Studies rated as at critical risk of bias in any domain will be excluded from the reporting of results, synthesis, and conclusion; however, study details will be included under '*Characteristics of included studies*'.

Acupressure

Systematic reviews

The methodological quality of included systematic reviews will be assessed using the AMSTAR-2 quality assessment checklist (45). The AMSTAR-2 consists of 16 domain questions (see Appendix C) that are answered as 'yes', 'no', or 'partial yes'; with a 'yes' answer denoting a positive result. Any notable strengths or limitations of the SR (in reference to the relevant AMSTAR-2 domains) will be reported. If the SR is broader in scope than the clinical question posed in this review (i.e. includes other interventions or NRSIs not eligible for inclusion), the overall quality of the systematic review will be assessed.

It is noted that the AMSTAR-2 leads to a judgement of methodological quality (or limitations) of a SR, not a judgement about risk of bias of the body of evidence included within the SR. Implications concerning relevant AMSTAR-2 items for the risk of bias of primary studies and assessing the certainty of evidence are discussed in Sections 3.3.14 and 3.3.16.Risk of bias of included studies

An independent assessment of the risk of bias of RCTs and pseudo-RCTs included within an eligible SR will not be performed. Instead, the risk of bias of these studies (or outcomes) will be as reported within the included SR; no further imputations will be made. A description of the quality assessment tool used to assess the studies will also be provided.

Where a study is included in multiple SRs, a crosscheck of the risk of bias assessment across SRs will be performed and any discrepancies will be reconciled based on available information. The best available and most comprehensive data will used when assessing the certainty of evidence (see Section 3.3.16). Footnotes will be used to document the source of all information.

In the absence of any risk of bias information for an individual study or when appropriate risk of bias information is not available (e.g. the SR reports risk of bias for the overall study and not at the outcome-level, or the SR does not use an appropriate tool to assess risk of bias), inferences about risk of bias when assessing the certainty of evidence will be made as described in Section 3.3.16.

3.3.7 Risk of bias assessment process

The risk of bias for each included study or SR will be assessed by one reviewer. The lead reviewer will then check and confirm all assessments made. Disagreements will be resolved by discussion, with advice sought from a third reviewer if agreement cannot be reached.

For each study or review, we will report our judgements for each domain or item and provide a rationale for the judgement with supporting information (see Appendix C). Overall risk of bias judgements will be described in the '*Characteristics of included studies'* table.

To ensure consistency among reviewers, pre-testing of risk of bias assessments will be achieved by all reviewers completing assessments for two SRs, two RCTs and two NRSIs (using AMSTAR-2, RoBv2.0 and ROBINS-I, respectively). Studies will be selected to cover the breadth of the PICO and included study design features. The lead reviewer will inspect the forms to ensure consistency, and any differences will be resolved through discussion.

Shiatsu

For each outcome, we will report our judgement of risk of bias (e.g. low, moderate, high, critical, unclear) by domain and provide a rationale for the judgement with supporting information. Overall risk of bias judgements will be described in the *'Characteristics of included studies'* table.

Acupressure

For each item on the AMSTAR-2 checklist, we will answer 'yes', 'no', or 'partial yes'. Limitations for each SR (including a rationale for judgements with supporting information) will be described in the *'Characteristics of included studies'* table.

3.3.8 Measures of effect

Shiatsu

Dichotomous data will be presented as risk ratios (RR) with 95% confidence intervals and *p*-values. Continuous data will be reported as mean difference (MD) (along with the standard deviation (SD) and number of participants). Standardised mean difference (SMD) will be used when different scales are used to measure the same conceptual outcome (e.g. function). To ensure that all the scales point in the same direction of effect, data from one set of studies will be adjusted before standardisation by multiplying the mean value by -1 to be consistent with the other set of studies. Time-to-event data will be presented as hazard ratios (HR) and, if analyses of covariance have been used to adjust for baseline measures, the adjusted effect estimates will also be recorded.

To reduce effects of confounding, summary statistics from NRSIs will preferably be reported as adjusted effect estimates (e.g. adjusted odds ratios (OR) from logistic regression). The variables that have been used for adjustment will be recorded.

As there are a broad range of populations eligible for inclusion in the review, it is not possible to prespecify the minimal clinically important differences for each outcome. However, where possible, the minimal clinically important difference will be sourced from published reports or will be guided by advice from the NTWC.

Acupressure

For SRs that do not report a meta-analysis for a relevant comparison because it is not appropriate to do so, effect estimates for the primary studies will be reported (where possible) as described above for Shiatsu. If appropriate to meta-analyse, analysis will be performed as per Section 3.3.11.

3.3.9 Unit-of-analysis issues

No imputation for unit-of-analysis issues will be performed.

While intervention-related clustering is a possibility in this review, it is considered unlikely that included studies will have provided adequate details to enable clusters to be accounted for in their analyses. No adjustments will be made for intervention-related clustering using a statistical method. However, where such clusters are identified, this will be noted in the relevant part of the review along with discussion of the potential impacts of the clustering on the review findings.

SRs that have included studies with potential for unit-of-analysis issues (i.e. cluster-randomised trials, crossover trials, repeated observations) will be noted in the results table, noting how the SR dealt with the unit-of-analysis issues in their evidence synthesis (if at all). The implications of the unit-of-analysis issues will be considered when interpreting the evidence, with any important implications for interpreting results documented in footnotes to the summary of findings table.

Potential unit-of-analysis issues relating to inclusion of the same study (and same group of participants) in both the shiatsu and acupressure reviews (see Section 3.3.11) will also be considered when interpreting the evidence.

Cluster-randomised trial

To avoid a unit-of-analysis error in a cluster-randomised trial we will extract and report effect estimates from analyses undertaken by the trial authors. Information regarding the approach used to account for the cluster design will be recorded. If the study authors have not provided information relating to the method of adjustment (i.e. the estimate of the relative variability within and between clusters), the implications of the missing data will be considered when interpreting the evidence and will be discussed under '*Overall completeness and applicability of evidence*'.

Crossover trial

To avoid a unit-of-analysis error in a crossover trial, only data from the first period will be included in the analysis. Studies reporting paired analysis will be discussed separately, and the potential impact of selective reporting will be discussed under '*Overall completeness and applicability of evidence*'.

Repeated observations

To avoid a unit-of-analysis error in studies reporting reports from more than one timepoint, results from a single timepoint will be selected, and only data from that timepoint will be presented in the analysis. The timepoint selected will be based on that determined to the critical or important for decision-making as outlined in Section 3.1.4.

3.3.10 Studies with more than two intervention groups

If the included studies have multiple treatment groups, only single pairwise comparisons of the intervention with a control (i.e. placebo, no intervention or other) will be considered. If appropriate to combine like groups, we will combine to create a single pairwise comparison. The combining of summary statistics across groups will be as described in Chapter 6 of the Cochrane Handbook (49).

SRs that have included studies with multiple treatment groups, will be noted in the results table along with discussion of how the SR dealt with the multiple treatment groups in the evidence synthesis (e.g. combining like groups to create a single pairwise comparison, double counting the placebo group in a meta-analysis). The implications of the multiple treatment groups will be

considered when interpreting the evidence with any important implications for interpreting results documented in footnotes to the summary of findings table.

3.3.11 Meta-analysis

Synthesis will only be undertaken for studies that compare shiatsu or acupressure with 'control' (stratified into 'placebo' or 'no intervention'). Results data from studies comparing shiatsu with 'other' interventions (or pooled data from SRs comparing acupressure with 'other' interventions) will be extracted and presented in data tables (see Appendix D), but will not be synthesised further, except where requested by the NTWC. Pooled data from SRs (for acupressure) will include a list of studies contributing data. Where pooled data is not available, individuals study results will be extracted (see Section 3.3.8). These data will be presented as an 'evidence inventory' to provide a snapshot of the available evidence comparing shiatsu or acupressure with 'other' interventions.

The NTWC may request that data comparing shiatsu or acupressure with an 'other' intervention be synthesised, where:

1. at least two studies compare the effect of shiatsu or acupressure with the same active comparator, and the comparator is sufficiently homogenous across studies to support synthesis, and
2. at least two of these studies are at low or moderate risk of bias, and
3. the comparator represents an accepted, evidence-based 'gold standard' of care for the population in question.

Such cases will be identified by the NTWC through blinded discussions with the evidence reviewer at the data synthesis stage, or prior to provision of the first draft evaluation report.

Shiatsu

Data from RCTs

Data synthesis of RCTs will be performed using RevMan 5.4 and forest plots presented. Studies will be compared in appropriate groups according to population, intervention (or component thereof) and comparator. Within each comparison we will combine effect estimates across studies for each outcome using a random effects model to take into account expected differences between studies. Statistical heterogeneity will be assessed by visually inspecting the overlap of confidence intervals on the forest plots, formally testing for heterogeneity using the Chi^2 test (using a significance level of α=0.1), and quantifying heterogeneity using the I^2 statistic (50).

Data from NRSIs

NRSIs will be analysed separately from RCTs. Where appropriate, data synthesis of NRSIs will be performed using RevMan 5.4 and forest plots presented. Within each comparison we will combine effect estimates across studies for each outcome using a random effects model to take into account expected differences between studies. Statistical heterogeneity will be assessed by visually inspecting the overlap of confidence intervals on the forest plots, formally testing for heterogeneity using the Chi^2 test (using a significance level of α=0.1), and quantifying heterogeneity using the I^2 statistic (50).

Effect estimates will only be combined if the included NRSIs are judged to be at low to moderate risk of bias for that outcome (see Section 3.3.6) and are sufficiently homogenous to be combined. This means the PICO criteria of the NRSIs must be sufficiently similar and the study design features should be comparable.

Acupressure

Data from SRs

If there are several eligible SRs identified that evaluate the effectiveness of acupressure across the same PICO, preference will be given to extracting pooled results (where available) from the best available source (e.g. the most recent and comprehensive SR) based on the process outlined in Figure 1.

Where the selected meta-analysis result identifies all available studies across the breadth of the PICO (i.e. all eligible comparisons and outcomes for a population or condition), pooled data from the selected SR will be presented with no further data synthesis; that is, summary effect estimates (95% confidence intervals, *p*-values) will be extracted as reported by the SR authors. The effect estimates of the primary studies will not be extracted, however the individual studies contributing data will be recorded. The meta-analysis model fitted, number of included studies, and any reported measures of heterogeneity will be included (e.g. I^2 statistic and associated *p*-value). If available, the certainty of evidence (GRADE) (and any sensitivity analysis) will also be recorded.

Where studies identified for inclusion in the shiatsu component of this review are also included within a meta-analysis (for acupressure), they will be excluded from the acupressure component by removal of the appropriate data from the meta-analysis (where possible). This is to avoid duplicate reporting of results. If it is not possible to remove the data from the meta-analysis, then the implications for reporting bias (e.g. unit of analysis issues associated with duplicate reporting within the review) will be discussed under 'Potential biases in the review process' and any indirectness within the SR (for acupressure) considered during the GRADE assessment.

Where the selected SR result does not include all eligible studies for a given comparison and outcome (or includes studies that are ineligible for inclusion), the meta-analysis reported by the selected SR will be updated (or re-analysed). Data from all eligible studies will be included, and any ineligible studies removed (where possible). The decision to re-analyse data will be determined by whether:

- PICO characteristics of any additional studies are judged to be sufficiently similar (based on comparisons relevant to the Overview question, rather than individual SR questions),
- required summary statistics are available (or able to be calculated) for that study,
- the SR presents sufficient data to facilitate the addition of eligible studies or removal of any studies ineligible for inclusion in this Overview, and
- the inclusion of results from the additional study or studies are likely to change the direction of effect (i.e. where the direction of effect is inconsistent with the pooled estimate of effect).

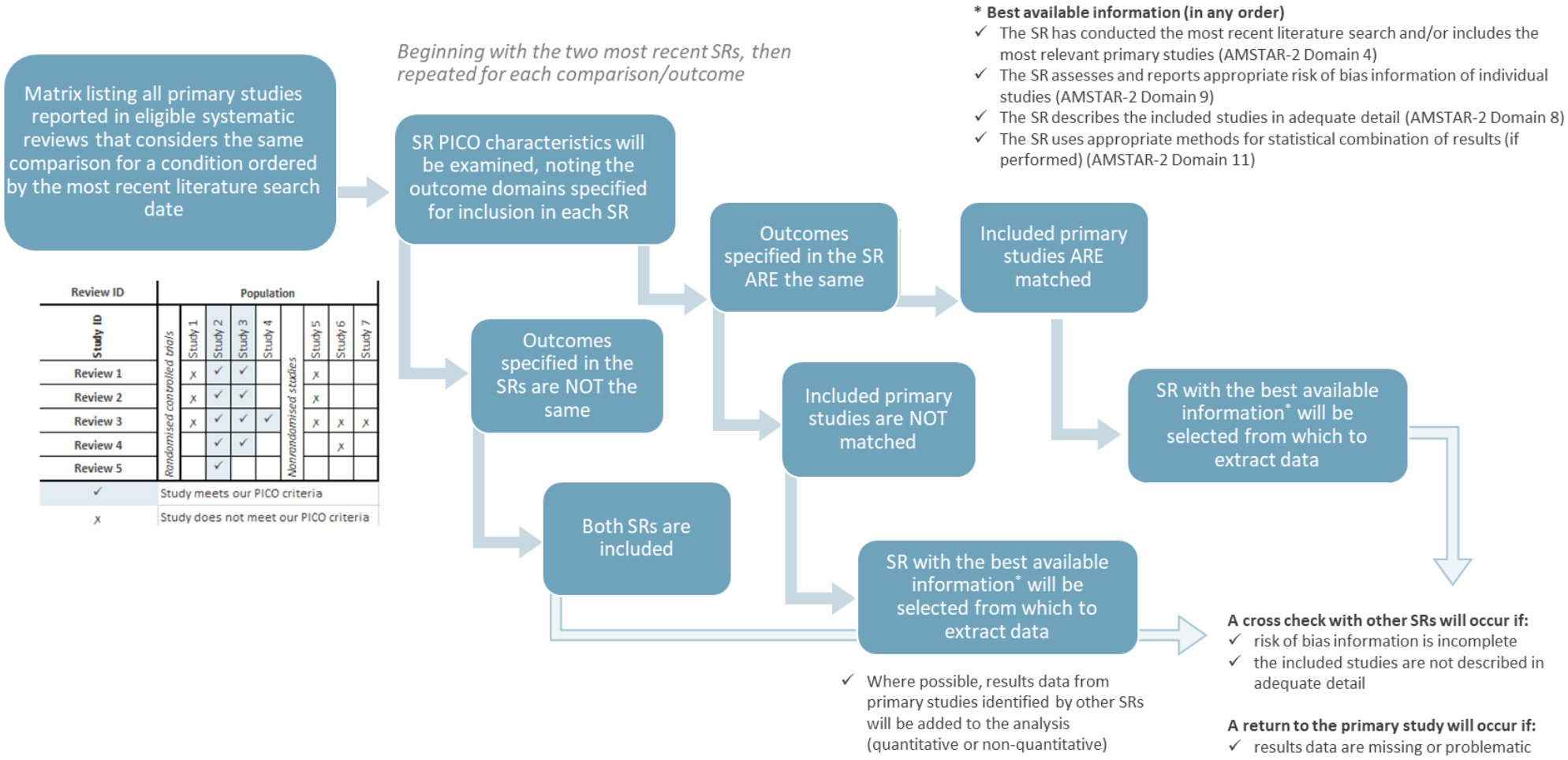
Where a meta-analysis of an eligible SR is found to include an ineligible study (e.g., includes other interventions or study designs not eligible for inclusion in this review), re-analysis will involve removal of the ineligible data from the meta-analysis (where possible). If it is not possible to remove the data from the meta-analysis, then the implications for indirectness will be considered during the GRADE assessment (see Section 3.3.16)

If, for a comparison, there is a mix of quantitative and qualitative data that is unable to be quantitatively synthesised (e.g. due to incomplete data or missing information), then a structured summary of the results will be presented (see Section 3.3.12).

If the best-available SR does not report a meta-analysis for a relevant comparison, and it is appropriate to do so, data synthesis will be performed using RevMan 5.4 and forest plots presented (see Section 3.3.8). Within each comparison we will combine effect estimates across studies for each outcome using a random effects model to take into account expected differences between studies. Statistical heterogeneity will be assessed by visually inspecting the overlap of confidence intervals on the forest plots, formally testing for heterogeneity using the Chi^2 test (using a significance level of α=0.1), and quantify heterogeneity using the I^2 statistic (50).

For SRs where the meta-analysis or primary study results are incompletely reported (e.g. no effect estimate is reported, but the direction of effect is reported along with a *p*-value), we will report the available information (see Section 3.3.8). If the reported information allows for calculation of effect estimates or imputation of missing statistics (e.g. SD), we will perform the calculations as described in Chapter 6 of the Cochrane Handbook (49).

Figure 1 Framework for selecting the SR from which to extract data, for any given comparison and outcome for acupressure



3.3.12 Summary and synthesis when meta-analysis is not possible

The evidence review will provide a structured summary of the results for each condition identified, that will include risk of bias assessments (of primary studies), and other intervention characteristics, in tables structured by intervention (shiatsu [further delineated to multi- or single components], or acupressure), comparator ('placebo', 'control', or 'other' intervention), outcome domain, study design (ordered and grouped by evidence from SRs, RCTs, NRSIs), then risk of bias and study size. Where possible, a visual representation of the results of included studies will be presented in a forest plot (without a summary estimate) grouped by study design features and risk of bias.

Shiatsu

The narrative summary will include a brief description of the condition and studies identified (including study design features, size, and population demographics). This will be followed by a summary of results grouped by intervention, comparator, and outcome domain. For shiatsu, results from each primary study will be reported, with the range and magnitude of observed effects noted. For studies where the results are incompletely reported (i.e. no effect estimate is reported, but the direction of effect is reported along with a *p*-value), we will report the available information.

To describe an overall effect across multiple studies for each outcome (for studies comparing shiatsu with control [placebo, or no intervention] only), we will use a simple vote count based on direction of effect (e.g. X/Y studies reported an effect favouring the intervention for the outcome Z and A/Y studies reported no difference between groups (total participants, range of effect sizes [or other values] reported). We were unable to perform a meta-analysis for this outcome because [add in reasons]..). Studies that are judged to be at low risk of bias and are powered to demonstrate an effect will take precedence over studies that are underpowered and or judged to be at moderate or high risk of bias (i.e., will be called out and highlighted in the text). Any important differences in study size or design features that may influence the interpretation of results will be considered and discussed in the text.

Qualitative descriptors describing the size of the effect (small, large etc.) will be used only where appropriate and will be based on the smallest difference that patients perceive as beneficial (or detrimental) for that outcome.

Acupressure

The narrative summary will include a brief description of the condition and reviews identified (including review criteria for inclusion or exclusion of studies, population demographics and other key features). Any notable weaknesses within a review, or inconsistency across reviews will be recorded. This will be followed by a summary of results grouped by intervention, comparator and outcome domain.

Details regarding the number of studies and number of participants that inform the data will be included, with a brief statement regarding any overlap of primary studies provided. Any important differences in review criteria or in control group risks that may influence the interpretation of results will be considered and discussed in the text.

If there are several eligible SRs identified that evaluate the effectiveness of acupressure across the same PICO, results will be reported from the selected SR based on pre-specified criteria, as outlined in Figure 1. In the absence of supplementary quantitative data, results from additional studies identified in other SRs will be described, with the range and distribution of observed effects noted.

To describe the overall effect, a simple vote count based on direction of effect will be used (e.g. XX RCTs were identified in the SR by YY, who reported a pooled effect favouring the intervention for outcome ZZ (MH fixed effects; effect size; 95%CI; p-value, I^2; GRADE). One additional RCT was identified (by SR) that also reported an effect favouring the intervention, however we were unable to add this to the quantitative synthesis because [add in reasons]…). Any important features of the additional studies that may influence the interpretation of results (e.g. risk of bias. study size, power) will be called out and highlighted in the text.

3.3.13 Risk of reporting bias across studies

Given the size and breadth of this review, it is expected that a broad approach to data synthesis will transpire. This means, summary estimates will include an outcome domain (e.g. pain) measured at a rough time point (e.g. latest timepoint reported), using any instrument. This will increase the number of studies that will be eligible for inclusion in a summary estimate. As noted in Section 3.3.5, the implications for missing data within studies will be considered when interpreting the evidence. Similarly, judgements regarding missing data across studies will be made based on available information (e.g. from '*Studies awaiting classification*' etc.) and discussed under '*Overall completeness and applicability of evidence*', noting that approaches for assessing bias due to missing studies (such as additional searching of clinical trial registers, grey literature, or other reports) will not be performed.

Shiatsu

If more than 10 RCTs are included for a particular PICO, funnel plots (of effect estimates against their standard errors) will be generated using RevMan 5.4 so as to determine possible reporting biases. If, after visual inspection of the funnel plot there is evidence of asymmetry (suggesting small-study effects or missing results), a brief statement about the potential impact on the overall conclusions of the evidence review will be included under the relevant sections of the review (including the *'Overall completeness and applicability of evidence'*). Other possible reasons for funnel plot asymmetry will also be considered at this time (e.g. poor methodological quality, true heterogeneity, chance) (51). No additional statistical analysis for testing for small-study effects will be conducted.

Acupressure

Reporting biases identified within SRs will be as described by the review authors. Any variability across reviews and any important flaws in individual reviews will be discussed under *'Quality of the evidence'*.

3.3.14 Addressing risk of bias

Shiatsu

All RCTs will be included in the review, regardless of judgements made regarding risk of bias. A description of the risk of bias in individual domains will be presented along with the estimated effect estimated using the risk of bias visualisation tool (52). To examine the impact of risk of bias, a sensitivity analysis will be conducted, with studies judged to be at high risk of bias to be removed from the analysis. The impact of this change will be noted and discussed under '*Quality of evidence*'.

NRSIs rated as being at critical risk of bias across one or more domain will not be included in the reporting of results, synthesis, and conclusion. A brief statement about the impact of the exclusion of NRSIs with critical risk of bias on the overall conclusions of the evidence review will be included

under the relevant sections of the review (including the *'Overall completeness and applicability of evidence*').

Acupressure

All eligible SRs will be included in the review, regardless of judgements made regarding methodological quality, noting that:

- methodological flaws in an SR do not reflect the risk of bias at the primary study level, which is the level at which results are synthesised, and
- the framework in Figure 1 aims to preferentially report results from SRs with fewer methodological limitations for any given comparison and/or PICO.

No formal analysis specifically addressing the risk of bias at the SR-level will be conducted; however, where there are concerns related to the methodological quality of a SR (e.g. due to concerns with study eligibility criteria, methods used to identify, select, or appraise studies, or concerns with interpretation of findings) we will attempt to mitigate the potential bias by cross-checking data across SRs, re-analysing and re-interpreting results, or returning to the primary study to check data. A brief statement about the impact of any changes made to the evidence reported by a SR (e.g., removal of a study due to inappropriate inclusion, change to the risk of bias assessment for that study, update on the data reported by the SR because they reported an incorrect number) on the overall conclusions of the evidence review will be included under the relevant sections of the review (including the *'Overall completeness and applicability of evidence*').

Judgements regarding the risk of bias of primary studies will be based on that reported in the SRs. If sensitivity analyses have been reported (e.g. removal of studies judged to be at high risk of bias), these will be considered as part of the GRADE assessment for that result (see Section 3.3.16).

3.3.15 Subgroup analyses

We do not plan to undertake pre-specified subgroup analyses of subsets of participants within studies. For shiatsu, if there is inconsistency between effect estimates, subgroup analysis will be used to explore possible sources of heterogeneity relating to delivery of the intervention (e.g. intensity, mode of delivery, who delivers). A standard test for heterogeneity across the subgroups will be reported.

For acupressure, sources of inconsistency between SRs for the same PICO will be considered by exploring eligibility criteria of the reviews as well as PICO characteristics within the included studies to provide a hypothesis for future reviews.

3.3.16 Certainty of the evidence

Across each population, we will assess the certainty of the evidence for each outcome using the GRADE approach (6). Evidence for shiatsu and acupressure will be evaluated separately, as will evidence from RCTs and NRSIs (shiatsu only). Further, only evidence comparing shiatsu (or acupressure) with 'placebo' and 'no intervention' will be presented (in separate 'Summary of Findings' tables).

The GRADE process provides a framework for determining the certainty of the evidence and is based on consideration of the following five factors:

- *Risk of bias.* Based on the summary assessment of bias across studies for each outcome reported for a comparison (53).

- *Inconsistency.* Based on heterogeneity in the observed intervention effects across studies that suggests important differences in the effect of the intervention and whether this can be explained (54).
- *Imprecision.* Based on interpretation of the upper and lower confidence limits in relation to a clinically important threshold (i.e., the confidence interval includes both appreciable benefit and harm); and whether the optimal information size has been reached (i.e., the total number of patients meets the required sample size for a sufficiently powered individual study). In the absence of a published clinically important threshold a rough guide will be used (i.e., a 25% relative risk reduction or increase) (55).
- *Indirectness.* Based on important differences between the review questions and the characteristics of included studies that may lead to important differences in the intervention effects (56).
- *Publication bias.* Based on the extent to which the evidence is available. Publication bias would be suspected when the evidence is limited to a small number of small trials (57).

The certainty of evidence will be categorised as follows:

- High (⨁⨁⨁⨁): further research is very unlikely to change the confidence in the estimate of effect.
- Moderate (⨁⨁⨁⊖): further research is likely to have an important impact in the confidence in the estimate of effect.
- Low (⨁⨁⊖⊖): further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- Very low (⨁⊖⊖⊖): any estimate of effect is very uncertain.

For each domain, a judgement will be made about whether there are serious, very serious or no concerns; resulting in an overall GRADE describing the certainty of evidence for each outcome. Footnotes will be used to record judgements made about downgrading (or upgrading) the evidence (see 'Summary of findings' table). Scoring of the certainty of the evidence will begin as 'high' for RCTs and NRSIs (score=4), which can be downgraded by –1 for each domain with serious concerns or –2 for very serious concerns (6, 58).

The certainty of evidence can also be upgraded in certain circumstances. Factors that will be considered for upgrading the evidence include the following:

- *Large magnitude of an effect.* When large or very large effect estimates are observed, and there is more confidence in the results (sufficient number of events to be precise).
- *Dose-response gradient.* When there is a clear relation between outcome and increasing exposure levels.
- *Effect of plausible residual confounding.* If there are clear factors that have likely led to an under-estimate of the true effect such as unmeasured or unknown determinants in the adjusted analysis that are likely to be distributed unequally between intervention and control groups (6).

Acupressure

Where the selected SR has reported a GRADE certainty of evidence for a PICO, we will replicate that information in the Overview, adjusting for any important differences between the eligibility criteria

of the SR and the Overview (i.e., we will assess indirectness of the evidence in relation to the overview PICOs). If the PICO from the selected SR has been reanalysed through the removal of studies that do not meet the review PICO (see Section 3.3.11) the GRADE assessment (if available) will be revised to reflect the totality of the evidence. If the PICO from the selected SR has been updated with supplementary data (see Section 3.3.11 Data from SRs) the GRADE assessment (if available) will be revised to reflect the totality of the evidence. If the selected SR has not evaluated the certainty of the evidence using GRADE, a *de novo* GRADE assessment will be performed.

In the absence of enough information from the selected SR (or other SRs) to inform our judgements for a certain GRADE domain (e.g. risk of bias, imprecision), the certainty of the evidence will be determined by the following system (adapted from the *NHMRC Report on the Evidence: promoting social and emotional development and wellbeing of infants in pregnancy and the first year of life*) (59):

- *Where information regarding one domain is missing and there are no apparent concerns or reasons to downgrade:* GRADE certainty of the evidence will be presented as a range including the rating based on the available information and one level lower to account for uncertainty in the rating (e.g. an outcome has serious concerns due to risk of bias, but there is no information on imprecision, then the overall certainty of evidence will be judged to be 'moderate to low quality').
- *Where information regarding one domain is missing and there is some indication that the factor with missing information should be downgraded:* the overall certainty of evidence will be downgraded based on the available information, and presented with 'assumed' in brackets. Footnotes will be used to record assumptions judgements made about downgrading the evidence in the absence of information.
- *Where information regarding two features is not reported:* the overall quality of the evidence will be judged to be 'unclear'.

Any weaknesses in the reporting of included SRs that may impact the finding of the Overview will be considered and discussed under *'Overall completeness and applicability of evidence*' and '*Potential biases in the overview process'*.

3.3.17 'Summary of findings' tables

For each population, findings for the critical and important outcomes (see Section 3.1.4), will be reported in 'Summary of Findings' tables that will be prepared using the GRADEpro GDT software (www.gradepro.org). The findings for shiatsu (from RCTs and NRSIs) and for acupressure (from SRs) will be presented separately. Estimates of treatment effects for each outcome will be reported as absolute and relative risks (or standardised means). In the absence of quantitative data, a narrative synthesis will be provided (see Section 3.3.12). All critical and important outcomes will be reported, regardless of whether the findings demonstrate a clinically meaningful change.

The 'Summary of Findings' tables will provide a summary of each of the included critical and important outcomes and the certainty of evidence rating for each outcome in a quick and accessible

format (6). As part of the summary of findings table, an evidence statement pertaining to each outcome will be included. This statement will be guided by the following format:

The practice of shiatsu / acupressure in [population] [is suggested to, may, results] in [little to no effect, reduce, increase, promote etc.] on [outcome] compared with [placebo or control].

A technical report that presents, in detail, the evidence base for each research question by outcome will be developed and will include the following:

- the methodology used to identify the evidence base (documented systematic literature search, inclusion and exclusion criteria described),
- the characteristics of the included studies, including data extraction and risk of bias forms, and
- detailed results, presented by population and outcome, containing complete information about the evidence assessment.

Contributions of authors

MJ wrote and developed the draft Research Protocol with contributions in writing sections, providing comment and proofreading final drafts from SA, SB, AM and AS. The search strategy was developed and tested by MJ and SA. AM and AS advised on the screening and data extraction process.

NTREAP and NTWC provided expert advice, especially in relation to intervention, study design and eligibility criteria. Cochrane Australia conducted a methodological review of the draft protocol.

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 NTWC members are lodged with the NHMRC and are available online.

This work is funded by the National Health and Medical Research Council (NHMRC) under Official Order 2019-20P026.

References

1. Australian Government Department of Health. The 2015 Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies 2019. Available from: https://www.health.gov.au/internet/main/publishing.nsf/Content/phi-natural-therapies.
2. Higgins J, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, et al. Methodological Expectations of Cochrane Intervention Reviews. London: Cochrane; 2019. Available from: https://community.cochrane.org/mecir-manual.
3. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4:1.
4. KP Health. Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies. Part A - Overview report for shiatsu 2014.
5. KP Health. Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies. Part B - Consideration of stakeholder submissions for shiatsu 2014.
6. 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; (Updated October 2013). Available from: https://gdt.gradepro.org/app/handbook/handbook.html.
7. Strapps E, Hunter J, Douglas D, Spence P. Practicing shiatsu in Australia: A workforce survey. Advances in Integrative Medicine. 2019.
8. Beresford-Cooke C. Shiatsu: Theory and practice. 3 ed. Edinburgh: Elsevier Science Ltd; 2016.
9. Stefanini P. 16 - Ki in shiatsu. In: Mayor D, Micozzi MS, editors. Energy Medicine East and West. Edinburgh: Churchill Livingstone; 2011. p. 211-21.
10. Lee EJ, Frazier SK. The efficacy of acupressure for symptom management: a systematic review. J Pain Symptom Manage. 2011;42(4):589-603.
11. Stevensen C. Shiatsu. Complement Ther Nurs Midwifery. 1997;3(6):168-70.
12. Shiatsu Therapy Association of Australia. About Us: Shiatsu Therapy Association of Australia; 2020 [4 January 2020]. Available from: https://www.staa.org.au/About.
13. CHAPTER 4 - Shiatsu Concepts. In: Anderson SK, editor. The Practice of Shiatsu. Saint Louis: Mosby; 2008. p. 72-83.
14. Robinson N, Lorenc A, Liao X. The evidence for Shiatsu: a systematic review of Shiatsu and acupressure. BMC complementary and alternative medicine. 2011;11:88.
15. Shiatsu Therapy Association of Australia. Styles of Shiatsu: Shiatsu Therapy Association of Australia; 2020 [4 January 2020]. Available from: https://www.staa.org.au/Styles-of-Shiatsu.
16. World Health Organization. Regional Office for the Western P. WHO international standard terminologies on traditional medicine in the Western Pacific Region. Manila: WHO Regional Office for the Western Pacific. Availabe from http://apps.who.int/iris/ 2007.
17. Cabo F, Baskwill A, Aguaristi I, Christophe-Tchakaloff S, Guichard JP. Shiatsu and Acupressure: Two Different and Distinct Techniques. Int J Ther Massage Bodywork. 2018;11(2):4-10.
18. Chirali IZ. 6 - Complementary and Alternative Medicine (CAM) Therapies that Can Safely Introduce Cupping to their Treatment Protocol. In: Chirali IZ, editor. Traditional Chinese Medicine Cupping Therapy (Third Edition): Churchill Livingstone; 2014. p. 72-8.
19. Ernst E. Massage therapy for low back pain: a systematic review. J Pain Symptom Manage. 1999;17(1):65-9.
20. Collected Reports of The Shiatsu Therapy Research Lab 1998-2012. Tokyo, Japan: The Japan Shiatsu College; 2013. Available from: http://e.shiatsu.ac.jp/research/.
21. Collected Reports of The Shiatsu Therapy Research Lab 2013-2017. Tokyo, Japan: The Japan Shiatsu College; 2018. Available from: http://e.shiatsu.ac.jp/research/.
22. Grandin T. Calming effects of deep touch pressure in patients with autistic disorder, college students, and animals. Journal of child and adolescent psychopharmacology. 1992;2(1):63-72.
23. Uvnas-Moberg K, Handlin L, Petersson M. Self-soothing behaviors with particular reference to oxytocin release induced by non-noxious sensory stimulation. Front Psychol. 2014;5:1529.
24. Goats GC. Massage–the scientific basis of an ancient art: Part 2. Physiological and therapeutic effects. Br J Sports Med 1994;28:153-6.
25. Sagar S, Dryden, T., Wong, K.,. Massage therapy for cancer patients: a reciprocal relationship between body and mind. . Curr Oncol. 2007;14:45-56.
26. Imamura M, Furlan AD, Dryden T, Irvin EL. CHAPTER 16 - Massage Therapy. In: Dagenais S, Haldeman S, editors. Evidence-Based Management of Low Back Pain. Saint Louis: Mosby; 2012. p. 216-28.

27. Long AF. The potential of complementary and alternative medicine in promoting well-being and critical health literacy: a prospective, observational study of shiatsu. BMC complementary and alternative medicine. 2009;9:19.
28. Armour M, Ee CC, Hao J, Wilson TM, Yao SS, Smith CA. Acupuncture and acupressure for premenstrual syndrome. Cochrane Database Syst Rev. 2018;8:CD005290.
29. Smith CA, Armour M, Dahlen HG. Acupuncture or acupressure for induction of labour. Cochrane Database Syst Rev. 2017;10:CD002962.
30. Smith CA, Collins CT, Crowther CA, Levett KM. Acupuncture or acupressure for pain management in labour. Cochrane Database Syst Rev. 2011(7):CD009232.
31. Sterne J, Hernán M, McAleenan A, Reeves B, Higgins J. Chapter 25: Assessing risk of bias in a non-randomized study. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 6,0 [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
32. Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L. Chapter 3: Systematic reviews of effectiveness. In: Aromataris E, Munn Z, editors. Joanna Briggs Institute Reviewer's Manual: The Joanna Briggs Institute; 2017.
33. Review Manager (RevMan) [Computer program]. Version 5.4. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration; 2020.
34. Higgins JPT, Eldridge S, Li T. Chapter 23: Including variants on randomized trials. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 6,0 (updated July 2019). Cochrane2019.
35. Merlin T, Weston A, Tooher R. Extending an evidence hierarchy to include topics other than treatment: revising the Australian 'levels of evidence'. BMC Med Res Methodol. 2009;9:34.
36. National Health and Medical Research Council. NHMRC Additional levels of evidence and grades for recommendations for developers of guidelines. 2009. Available from: https://www.mja.com.au/sites/default/files/NHMRC.levels.of.evidence.2008-09.pdf.
37. Reeves B, Deeks J, Higgins J, Shea B, Tugwell P, Wells G. Chapter 24: Including non-randomized studies on intervention effects. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 6,0 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
38. Reeves B, Higgins J, Ramsay C, Shea B, Tugwell P, Wells G. An introduction to methodological issues when including non-randomised studies in systematic reviews on the effects of interventions. Res Synth Methods. 2013;4(1):1-11.
39. Schunemann HJ, Tugwell P, Reeves BC, Akl EA, Santesso N, Spencer FA, et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. Res Synth Methods. 2013;4(1):49-62.
40. National Health and Medical Research Council. How to use the evidence: assessment and application of scientific evidence. 2000. Available from: https://www.nhmrc.gov.au/about-us/publications/how-use-evidence.
41. Cochrane Childhood Cancer. 2020 [Available from: https://childhoodcancer.cochrane.org/non-randomised-controlled-study-nrs-designs.
42. Australian Institute of Health and Welfare. Australia's health 2016. Canberra; 2016.
43. Brim RL, Miller FG. The potential benefit of the placebo effect in sham-controlled trials: implications for risk-benefit assessments and informed consent. J Med Ethics. 2013;39(11):703-7.
44. National Blood Authority. Patient Blood Management Guidelines: Module 6 Neonatal and Paediatrics. Technical report - Volume 2. Canberra, Australia: National Blood Authority; 2016. Available from: https://www.blood.gov.au/pbm-module-6.
45. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.
46. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
47. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 6,0. (updated July 2019): Cochrane 2019.
48. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.
49. Higgins JPT, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 6,0 (updated July 2019).
50. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.

51. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002.
52. McGuinness LA. robvis: An R package and web application for visualising risk-of-bias assessments. 2019.
53. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol. 2011;64(4):407-15.
54. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. J Clin Epidemiol. 2011;64(12):1294-302.
55. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol. 2011;64(12):1283-93.
56. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol. 2011;64(12):1303-10.
57. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. J Clin Epidemiol. 2011;64(12):1277-82.
58. Schunemann HJ, Cuello C, Akl EA, Mustafa RA, Meerpohl JJ, Thayer K, et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. J Clin Epidemiol. 2019;111:105-14.
59. Australian Research Centre for the Health of Women and Babies (ARCH). Evaluation of evidence on the effectiveness of interventions for caregiving practices and behaviours for optimal social and emotional development of infants: an overview of systematic reviews. https://www.nhmrc.gov.au/file/5391/download?token=k6mJopkU The University of Adelaide; 2017.

Appendix A – Literature search strategy

Concept: Study design limits (SR, RCT, NRSI, not animals)

1. exp meta analysis/ or meta analysis.mp. or exp systematic review/ or systematic review.mp. or pooled analysis.mp. or ((exp review/ or review.mp.) and (systemat* or pool*).mp.)
2. exp comparative study/ or comparative study.mp. or exp clinical trial/ or clinical trial.mp. or randomized controlled trial.mp. or randomi?ed controlled trial.mp. or exp randomized controlled trial/ or exp randomization/ or randomization.mp. or randomi?ation.mp. or exp single blind procedure/ or single blind procedure.mp. or exp double blind procedure/ or double blind procedure.mp. or exp triple blind procedure/ or triple blind procedure.mp. or exp crossover procedure/ or crossover procedure.mp. or exp placebo/ or placebo*.mp. or random*.mp. or rct.mp. or single blind.mp. or single blinded.mp. or double blind.mp. or double blinded.mp. or treble blind.mp. or triple blind.mp. or triple blinded.mp. or exp prospective study/ or prospective study.mp.
3. exp clinical study/ or exp case control study/ or exp family study/ or exp longitudinal study/ or exp retrospective study/ or exp cohort analysis/ or (cohort adj1 stud*).mp. or (case control adj1 stud*).mp. or (exp prospective study/ not randomi?ed controlled trials.mp.) or (follow up adj1 stud*).mp. or (observational adj1 stud*).mp. or (epidemiologic* adj1 stud*).mp. or (cross sectional adj1 stud*).mp.
4. case report/
5. (editorial or letter or comment or historical article).pt.
6. (animals/ or nonhuman/) not humans/
7. 4 or 5 or 6

Concept: Acupressure

8. exp acupressure/
9. Digitopression.ti,ab.
10. Digitopuncture.ti,ab.
11. Acupresion.ti,ab.
12. Acupressure.ti,ab.
13. Acupression.ti,ab.
14. Acupressao.ti,ab.
15. acupress*.ti,ab.
16. or/8-15

Concept: shiatsu or amna

17. exp shiatsu/
18. (Japanese adj2 massage).ti,ab.
19. Chih Ya.ti,ab.
20. Zhi Ya.ti,ab.
21. Shiat?u.ti,ab.
23. Shiatsu.ti,ab.
23. shiatzu.ti,ab.
24. anma.ti,ab.
25. amna.ti,ab.
26. amma.ti,ab.
27. gua gon.ti,ab.
28. gua sha.ti,ab.
29. or/17-28

Concept: acupoint

30. (acupoint and massage).ti,ab.
31. (acupoint adj3 pressure).ti,ab.
32. acupoint therap*.ti,ab.
33. (meridian* and (stretching or massage)).ti,ab.
34. (meridian* adj3 pressure).ti,ab.
35. (meridian* adj3 manipulat*).ti,ab.
36. or/30-35

Concept: SR of acupressure or acupoint

37. 1 AND (16 or 36)
38. 37 NOT 7

Concept: RCT of shiatsu or acupoint

39. 2 AND (29 or 36)
40. 39 NOT 7

Concept: NRSI of shiatsu or acupoint

41. 3 AND (29 or 36)
42. 41 NOT 7

The above strategy was developed in-house and will be adapted to suit EBSCO, the Cochrane Library and PubMed (limited to in-process citations and citations not indexed in MEDLINE).

Ovid syntax

Exp explodes controlled vocabulary term (i.e. includes all narrower terms in the hierarchy)
* denotes a term that has been searched as a major subject heading
/ denotes controlled vocabulary terms (EMTREE)
$ truncation character (unlimited truncation)
$*n* truncation limited to specified number (*n*) of characters (e.g. time$1 identifies time, timed, timer, times but not timetable)
* truncation character (unlimited truncation)
? substitutes any letter (e.g. oxidi?ed identifies oxidised and oxidized)
adj*n* search terms within a specified number (*n*) of words from each other in any order
.ti. limit to title field
.ti,ab. limit to title and abstract fields
.kw,ti,ab. limit to keyword, title and abstract field
.pt limit to publication type

CINAHL syntax

* truncation character (unlimited truncation)
# wildcard character will replace 1 or 0 characters (e.g. f#etus will retrieve fetus and foetus)
? wildcard character will replace one character (e.g. wom?n will retrieve women and woman)
MH - Search the exact CINAHL® subject heading; searches both major and minor headings
MH"heading"+ Search an exploded subheading
TI search title fields
AB search abstract fields
Nn – Proximity "near" operator will find a result if the terms are within a certain number (n) words of each other, regardless of the order in which they appear. (e.g. eating N5 disorders for results that contain eating disorders, as well as mental disorders and eating pathology.)
PT limit to publication type

PubMed syntax

The PubMed search will be restricted to records that are not indexed for MEDLINE (i.e. in-process citations and citations from journals (or parts of journals) that are not currently MEDLINE-indexed)

The search will comprise free-text terms only and replicates the free-text sets in the Embase search (converted from the Ovid syntax).

* truncation character (unlimited truncation)
[TI] limit to title field
[TIAB] limit to title and abstract fields
[EDAT] date citation added to PubMed
[SB] PubMed subset

AND pubmednotmedline[sb] will be added to the last line of search string

Appendix B – Screening criteria

A priori screening criteria are listed below. Items 1 through 8 will be considered and applied at abstract/title screening. All items will be considered and applied as appropriate at full text review (these studies will be listed in the technical report with reasons for exclusion):

Shiatsu

1. Duplicate citation
2. Nonhuman study
3. Intervention out of scope (not shiatsu or component thereof)
4. Population out of scope (healthy participants seeking general wellness)
5. Comparator out of scope
6. Outcome out of scope (patient-reported experiences, safety, quality and economic outcomes)
7. Publication type out of scope
    a. opinion piece/editorial/commentary
    b. not an intervention study examining effectiveness
8. Study design out of scope (specify)
    a. non-systematic review, Guidelines or HTA assessment
    b. SR of RCTs or NRSIs
    c. nonrandomised study
    d. case series or other
9. Duplicate citation submitted to the department (RCT/NRSIs already identified in this SR)
10. Publication not available in English [a]
11. Other (specify):
    a. duplicate data (multiple reports arising from the same study)
    b. superseded (more recent data from the primary study is available)
    c. withdrawn
    d. erratum
12. Relevant but additional followup needed (specify)[b]
    a. conference proceeding (data incomplete)
    b. ongoing study (results not available)
    c. no outcome of interest reported

Acupressure

1. Duplicate citation
2. Nonhuman study
3. Intervention out of scope (not acupressure)
4. Population out of scope (healthy participants seeking general wellness)
5. Comparator out of scope
6. Outcome out of scope (patient-reported experiences, safety, quality and economic outcomes)
7. Publication type out of scope
    a. opinion piece/editorial/commentary
    b. not an intervention study examining effectiveness
8. Study design out of scope (specify)
    a. non-systematic review

   b. RCT or pseudo RCT
   c. NRSI
   d. case series or other
9. Duplicate citation submitted to the department (RCT/NRSIs already identified in this SR)
10. Publication not available in English
11. Other (specify)
   a. duplicate data (multiple reports arising from the same study)
   b. superseded (SR has been updated or more recent data available)
   c. withdrawn
   d. erratum
12. Relevant but additional followup needed (specify) [b]
   a. conference proceeding (author contacted but no further information)
   b. SR protocol only (results not yet available)
   c. no outcome of interest reported

a. Screening of articles not published in English will be conducted as described in the Section 3.3.1 '*Studies published in languages other than English*'.
b. Articles tagged as relevant but additional followup needed are included but will not be incorporated in the evidence appraisal. These studies may be listed as '*Studies awaiting classification*', '*Ongoing*', or may be considered when developing conclusions about the '*Overall completeness and applicability of evidence*'.

Appendix C – Risk of bias forms

AMSTAR-2 (systematic reviews)

1. Did the research questions and inclusion criteria for the review include the components of PICO?			
	For Yes: ☐ Population ☐ Intervention ☐ Comparator groups ☐ Outcome	Optional (recommended) ☐ Timeframe for follow-up	☐ Yes ☐ No
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?			
	For Partial Yes: The authors state that they had a written protocol or guide that included ALL the following: ☐ review question(s) ☐ a search strategy ☐ inclusion/exclusion criteria ☐ a risk of bias assessment	For Yes: As for partial yes, plus the protocol should be registered and should also have specified: ☐ a meta-analysis/synthesis plan, if appropriate ☐ AND a plan for investigating causes of heterogeneity ☐ justification for any deviations from the protocol	☐ Yes ☐ Partial Yes ☐ No
3. Did the review authors explain their selection of the study designs for inclusion in the review?			
	For Yes, the review should satisfy ONE of the following: ☐ Explanation for including only RCTs ☐ OR Explanation for including only NRSI ☐ OR Explanation for including both RCTs and NRSI		☐ Yes ☐ No
4. Did the review authors use a comprehensive literature search strategy?			
	For Partial Yes (all the following): ☐ searched at least 2 databases (relevant to research question) ☐ provided key word and/or search strategy ☐ justified publication restrictions (e.g. language)	For Yes, should also have (all the following): ☐ searched the reference lists /bibliographies of included studies ☐ searched trial/study registries ☐ included/consulted content experts in the field ☐ where relevant, searched for grey literature ☐ conducted search within 24 months of completion of the review	☐ Yes ☐ Partial Yes ☐ No
5. Did the review authors perform study selection in duplicate?			
	For Yes, either ONE of the following: ☐ at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include ☐ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.		☐ Yes ☐ No
6. Did the review authors perform data extraction in duplicate?			
	For Yes, either ONE of the following: ☐ at least two reviewers achieved consensus on which data to extract from included studies ☐ OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.		☐ Yes ☐ No
7. Did the review authors provide a list of excluded studies and justify the exclusions?			
	For Partial Yes:	For Yes, must also have:	☐ Yes

	☐ provided a list of all potentially relevant studies that were read in full-text form but excluded from the review	☐ Justified the exclusion from the review of each potentially relevant study	☐ Partial Yes ☐ No
8. Did the review authors describe the included studies in adequate detail?			
	For Partial Yes (ALL the following): ☐ described populations ☐ described interventions ☐ described comparators ☐ described outcomes ☐ described research designs	For Yes, should also have ALL the following: ☐ described population in detail ☐ described intervention in detail (including doses where relevant) ☐ described comparator in detail (including doses where relevant) ☐ described study's setting ☐ timeframe for follow-up	☐ Yes ☐ Partial Yes ☐ No
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?			
	RCTs For Partial Yes, must have assessed RoB from ☐ unconcealed allocation, and ☐ lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)	For Yes, must also have assessed RoB from: ☐ allocation sequence that was not truly random, and ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome	☐ Yes ☐ Partial Yes ☐ No ☐ Includes only NRSI
	NRSI For Partial Yes, must have assessed RoB: ☐ from confounding, and ☐ from selection bias	For Yes, must also have assessed RoB: ☐ methods used to ascertain exposures and outcomes, and ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome	☐ Yes ☐ Partial Yes ☐ No ☐ Includes only RCTs
10. Did the review authors report on the sources of funding for the studies included in the review?			
	For Yes ☐ Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information, but it was not reported by study authors also qualifies		☐ Yes ☐ No
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?			
	RCTs For Yes: ☐ The authors justified combining the data in a meta-analysis ☐ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present. ☐ AND investigated the causes of any heterogeneity		☐ Yes ☐ No ☐ No meta-analysis conducted
	For NRSI For Yes: ☐ The authors justified combining the data in a meta-analysis ☐ AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present ☐ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available ☐ AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review		☐ Yes ☐ No ☐ No meta-analysis conducted

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?		
	For Yes: ☐ included only low risk of bias RCTs ☐ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.	☒ Yes ☐ No ☐ No meta-analysis conducted
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?		
	For Yes: ☐ included only low risk of bias RCTs ☐ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results	☐ Yes ☐ No
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?		
	For Yes: ☐ There was no significant heterogeneity in the results ☐ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review	☐ Yes ☐ No
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?		
	For Yes: ☐ performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias	☐ Yes ☐ No ☐ No meta-analysis conducted
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?		
	For Yes: ☐ The authors reported no competing interests OR ☐ The authors described their funding sources and how they managed potential conflicts of interest	☐ Yes ☐ No

Source: Amstar-2 Shea 2017 (45)

Cochrane RoB v2.0 (randomised controlled trials)

Study ID		
Domain	**Judgement**	**Description [a]**
Bias arising from the randomisation process	High risk Some concerns Low risk	Describe the method used to determine if: • the allocation sequence was random; • the allocation sequence was adequately concealed; • baseline differences between intervention groups suggest a problem with the randomisation process.
Bias due to deviations from intended interventions	High risk Some concerns Low risk	Describe the method used to conceal treatment allocation: • were participants aware of their assigned intervention during the trial? • were carers and people delivering the interventions aware of participants' assigned intervention during the trial? • consider the potential effect of deviations due to assignment and deviations due to adherence
Bias due to missing outcome data	High risk Some concerns Low risk	Describe the completeness of outcome data for each primary/secondary outcome, including whether • data for this outcome were available for all, or nearly all, participants randomised; • (if applicable) there was evidence that the result was not biased by missing outcome data; • (if applicable) missingness in the outcome was likely to depend on its true value (e.g. the proportions of missing outcome data, or reasons for missing outcome data, differ between intervention groups).
Bias in measurement of the outcome	High risk Some concerns Low risk	Describe whether: • the method of measuring the outcome was inappropriate; • measurement or ascertainment of the outcome could have differed between intervention groups; • outcome assessors were aware of the intervention received by study participants; • (if applicable) assessment of the outcome was likely to have been influenced by knowledge of intervention received.
Bias in selection of the reported result	High risk Some concerns Low risk	Describe whether: • the trial was analysed in accordance with a prespecified plan that was finalised before unblinded outcome data were available for analysis; • the numerical result being assessed is likely to have been selected, on the basis of the results, from multiple outcome measurements within the outcome domain; • the numerical result being assessed is likely to have been selected, on the basis of the results, from multiple analyses of the data.
Overall risk of bias		

Abbreviations:

Source: Adapted from Chapter 8 Cochrane handbook for systematic reviews of interventions (47).

Notes:

a. For the precise wording of signalling questions and guidance for answering each one, see the full risk of bias tool at www.riskofbias.info.

ROBINS-I (nonrandomised studies of interventions)

Study ID		
Domain	**Judgement**	**Comments**
Bias due to confounding	Low risk Moderate risk Serious risk Critical risk No information	Describe whether: • there is potential for confounding of the effect of intervention in this study; • the analysis is based on splitting participants' follow up time according to intervention received; • intervention discontinuations or switches were likely to be related to factors that are prognostic for the outcome; • the authors use an appropriate analysis method that controlled for all the important confounding domains; • the variables adjusted for are valid and reliable measures of the confounding domains; • the authors control for any post-intervention variables that could have been affected by the intervention; • the authors use an appropriate analysis method that adjusted for all the important confounding domains and for time-varying confounding; • confounding domains that were adjusted for measured validly and reliably by the variables available in this study;
Bias in selection of participants into the study	Low risk Moderate risk Serious risk Critical risk No information	Describe whether: • the selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention; • the start of follow up and start of intervention coincide for most participants; • adjustment techniques used are likely to correct for the presence of selection biases;
Bias in classification of interventions	Low risk Moderate risk Serious risk Critical risk No information	Describe whether: • the intervention groups are clearly defined; • the information used to define intervention groups is recorded at the start of the intervention; • classification of the intervention status has been affected by knowledge of the outcome or risk of the outcome;
Bias due to deviations from intended interventions	Low risk Moderate risk Serious risk Critical risk No information	Describe whether: • deviations from the intended intervention is beyond what would be expected in usual practice; • deviations from intended intervention is unbalanced between groups and likely to have affected the outcome; • important cointerventions balanced are across intervention groups; • intervention is implemented successfully for most participants; • study participants adhere to the assigned intervention regimen; • an appropriate analysis used to estimate the effect of starting and adhering to the intervention;
Bias due to missing data	Low risk Moderate risk Serious risk Critical risk No information	Describe whether: • outcome data is available for all, or nearly all, participants; • participants are excluded due to missing data on intervention status; • the proportion of participants and reasons for missing data are similar across interventions; • results were robust to the presence of missing data; • the predicted direction of bias is due is due to missing data.

Study ID		
Domain	**Judgement**	**Comments**
Bias in measurement of outcomes	Low risk Moderate risk Serious risk Critical risk No information	Describe whether: • the outcome measure has been influenced by knowledge of the intervention received; • outcome assessors are aware of the intervention received by study participants; • the methods of outcome assessment are comparable across intervention groups; • any systematic errors in measurement of the outcome related to intervention received.
Bias in selection of the reported result	Low risk Moderate risk Serious risk Critical risk No information	Describe whether the reported effect estimate likely to be selected, on the basis of the results: • multiple outcome measurements within the outcome domain; • multiple analyses of the intervention-outcome relationship; and • different subgroups.
Overall risk of bias		

Abbreviations:

Source: Adapted from Sterne 2019 (31)

Notes:

a. For the precise wording of signalling questions and guidance for answering each one, see the full risk of bias tool at www.riskofbias.info.

Appendix D – Data extraction forms

Characteristics of included studies

Systematic reviews

Review ID	**Author date**				
Review Title					
Review objective					
Author affiliations					
Source of funds					
Declared interests of the review authors					
Review method of analysis	e.g. Narrative, meta-analysis, Guidelines, HTA report				
Inclusion criteria					
Study design					
Population					
Intervention					
Comparator					
Other					
Exclusion criteria					
Study design					
Population					
Intervention					
Comparator					
Other					
Date of documented search (month/year)	(Include date range)				
Databases searched (list)	PubMed	Cochrane	Embase	etc.	
Were studies in a language other than English searched or included?	Yes	No	Not specified	Other comments	
Outcomes of Systematic Review (list)					
Primary					
Secondary					
Not specified					
Characteristics of studies included in the SR (Study ID, study design features, setting, other notable features)					
Study 1					
Study 2					
add rows as necessary					
Risk of bias of the included studies as reported in the systematic review (tool used, authors summary)					

Review ID	**Author date**		
Study 1	Tool used	Authors summary	
Study 2	Tool used	Authors summary	
Authors conclusions (key message)			
Studies meeting the inclusion criteria for this review (Study ID, no. of participants, RoB [as reported by the SR authors])			
Study 1	Dai 2007a	493	Low (other), high (selection bias), unclear (randomisation and outcome reporting)
Study 2			
Studies NOT meeting the inclusion criteria for this review (Study ID, no. of participants, Reason)			
Study 1	Shariati 2010	25	Pilot study to inform Shariati 2012 (patients included in Shariati 2012)
Study 2			

INTERNAL VALIDITY	
Overall methodological quality of the review (select from list)	e.g. Moderate More than one non-critical weakness – the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
Summary (descriptive)	e.g. The authors did not provide a full list of excluded studies or details relating to risk of bias assessments. Information regarding individual studies were limited. GRADE profiles were presented, and a comprehensive search strategy conducted.

Primary studies

Study ID	**Author date**				
Study design	RCT/NRSI		Features		
Affiliation/source of funds	Source of funding and conflicts of interest not declared				
Setting (single centre, multicentre, country/s)	Three hospitals	Tuscany region, Italy	Nursing home		Single provider of intervention
Enrolment period	Aug 2014 to Mar 2016				
Length of follow up (months)	12 months				
Description of population (# participants, age, comorbidities etc.)	N=	Elderly patients with osteoporosis at risk of falls			
Description of intervention/comparator (reported as per TIDIER checklist) (# of participants, # treatment session, session duration, program duration, co-interventions)					
Intervention	n=				
Comparator #1 (placebo)	n=				
Comparator #2 (control)	n=				
Comparator #3 (other)	n=				
Is instructor certified? (select from list)	Yes	No	Not specified	Additional comments	
Is comparator clearly inactive? (select from list)	Yes	No	Uncertain (seek advice)		
Outcomes	(list, description, measurement tool, timing)				
Primary #1	Pain	Numerical pain rating scale	0-100	higher score means worse pain	Short term
Primary #2					
Secondary #1					
Secondary #2					
Secondary #3					
add rows as needed					
Method of analysis					
Statistics	Descriptive, student t-test, regression				
Population analysed	ITT	PP	Other details		
Missing data	e.g. imputations, loss to follow up				

Abbreviations: ITT, intent-to-treat; PP, per protocol

Notes:

Characteristics of Studies awaiting classification

Study ID	**Author date**
Study description	Title or other descriptive text
Study design	RCT/NRSI
Participants	
Intervention	
Comparator	
Outcomes	
Notes	Article written in Korean with no English abstract - awaiting translation

Outcome data from included studies

Shiatsu - Primary studies (as reported by the study authors)

Study ID *RoB*	Population	Comparison	Outcome	Timing	Measured with	Measure details	No. participants (N)	[intervention] n/N (%) or mean (SD)	[comparator] n/N (%) or mean (SD)	Point estimate (95% CI)	*p*-value	Direction of effect
			[outcome #1]	3 months after randomisation	e.g. VAS scale 1-100	higher score means more pain				RR 1.00 [0.68, 1.48]	*X*	*No difference*
			[outcome #2]	> 3 months but less than 1 year from randomisation							*NR*	*Favours intervention*
			[outcome #3]	> 1 yr from randomisation								*Not reported*
			[outcome #4]	Not specified								
			[outcome #5]	Not specified								
			[outcome #6]	Not specified								
			[outcome #7]	Not specified								

Abbreviations: CI, confidence interval; hrs, hours; NR, not reported; OR, odds ratio; RoB, risk of bias; RR, relative risk; yr, year;

Acupressure – Table for data from systematic reviews and primary studies (as reported by the SR authors)

Review ID *AMSTAR-2*	**Population**	**Outcome**	**Measurement tool**	**Timing**	**Description**	**Included studies**	**No. participants (N)**	**[intervention] n/N (%) or mean (SD)**	**[comparator] n/N (%) or mean (SD)**	**Point estimate (95% CI)**	***p*-value**	**Direction of effect**	**Heterogeneity I^2 (*p*-value)[a]**	**Outcome RoB[b]**	**Certainty of evidence (GRADE)[b]**
Author date High	Chronic kidney disease	[outcome #1]	e.g. VAS scale 1-100	< 3 months	higher score means more pain	Study ID 1 Study ID 2 etc.	367			RR 1.23 [0.68, 1.48]	*p* = X	Favours intervention	77% < 0.0001 Considerable	summary across studies	High (⊕⊕⊕⊕)
		[outcome #2]	PSQI scale 0-21	> 3 months < 1 yr	higher score means better sleep quality							Favours comparator	Substantial		Moderate (⊕⊕⊕⊖)
		[outcome #3]		> 1 yr								Neutral	Moderate		Low (⊕⊕⊖⊖)
		[outcome #4]	Not specified									Not reported	Moderate		Very low (⊕⊖⊖⊖)
		[outcome #5]	Not specified										Mild		Not reported
		[outcome #6]	Not specified										No significant heterogeneity		High (⊕⊕⊕⊕)
		[outcome #7]	Not specified					Descriptive text if data not provided					Substantial		

Abbreviations: CI, confidence interval; hrs, hours; NR, not reported; OR, odds ratio; RoB, risk of bias; RR, relative risk; yr, year;

Notes:

a. Only applicable to Level I studies with formal meta-analysis. Heterogeneity defined as follows: (i) no significant heterogeneity if P_{het} >0.1 and I^2 <25%; (ii) mild heterogeneity if I^2 <25%; (iii) moderate heterogeneity if I^2 between 25–50%; (iv) substantial heterogeneity if I^2 >50% but <75%; (v) considerable heterogeneity if $I^2 \geq 75\%$.

b. Data as reported by the SR authors.