Project
Tai Chi 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

July 2020

Contents

Protocol information

Authors

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

[1] Health 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 NHMRC Natural Therapies Working Committee on 9 July 2020.

History

The National Health and Medical Research Council (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 Tai Chi. Eligible studies received from the Department's public 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 review the evidence regarding Tai Chi. 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 abbreviations

BRISA	Regional Base of Health Technology Assessment Reports of the Americas
CINAHL	Cumulative Index to Nursing and Allied Health Literature
COMET	Core Outcome Measures in Effectiveness Trials
GRADE	Grading of Recommendations Assessment, Development and Evaluation
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
ITT	Intent-to-treat
OR	Odds ratios
PAHO	Pan American Health Organization
PICO	Population, Intervention, Comparator, Outcome
PP	Per protocol
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	randomised controlled trial
RoB	Risk of bias
RR	Risk ratios
SD	Standard deviation
SR	Systematic review
TIDIER	Template for Intervention Description and Replication

1 Background

In 2015, a review of Tai Chi found very low quality evidence to suggest there may be some beneficial health effects from Tai Chi when compared to control in a limited number of conditions for a limited number of outcomes (see Section 1.4) (4, 5). The 2015 review was underpinned by an overview of systematic reviews that focused solely on Tai Chi and were published in the English language between 2008 and January 2014. Randomised controlled trials (RCTs) that were reported within included SRs and assessed Tai Chi delivered to treat any clinical condition were included, with outcomes selected according to predefined criteria. In this 2020 update, the evidence review will build upon the 2015 review but will not be limited by publication date and a broader range of study types will be eligible for inclusion (inclusive of pseudorandomised studies and, for certain populations and outcomes, nonrandomised studies of interventions [NRSIs]). The updated review will also include studies that assess Tai Chi delivered for primary prevention. Similar to the 2015 review, eligible comparisons will be Tai Chi versus control and Tai Chi 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, the Department's 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).

1.1 Description of condition and setting

Tai Chi is practiced for a range of reasons and is intended to improve general health and wellbeing.

Practicing Tai Chi is claimed to improve health outcomes for a variety of clinical and pre-clinical conditions, including:

- cardiovascular diseases, such as heart failure (7) and hypertension (8);
- pulmonary diseases, such as chronic obstructive pulmonary disease (9) and asthma (9);
- conditions associated with chronic pain, such as rheumatoid arthritis (10), osteoarthritis (11), peripheral neuropathy (12), fibromyalgia (13), migraine (14), and nonspecific low back pain (15);
- conditions associated with impaired physical functioning, such as multiple sclerosis (16), Parkinson's disease (17), or rehabilitation after stroke (18);
- conditions related to cancer survival (19-23);
- mental or behavioural disorders, such as depression (24), anxiety (20, 25);
- and conditions associated with ageing, such as osteoporosis (26), dementia (27), and mild cognitive impairment (28).

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

1.2 Description of intervention

Tai Chi is a mind-body movement practice developed from a traditional Chinese martial art 'Wushu', that combines slow physical movements with deep breathing and meditation (10). Practiced since at least the 16th Century, the integrated belief of a balance between mind, body and breath to achieve greater awareness, strength and a sense of well-being has evolved in agreement with Chinese philosophical principle of Yin-Yang (the balance of opposite, yet interdependent forces e.g. masculine/feminine, light/dark) (29). By nurturing, balancing and promoting the free flow of a person's life energy or vital force, termed in Chinese philosophy as 'qi', Tai Chi is believed to correct imbalances and interruptions to an individual's overall functioning and wellbeing. Also known as taijiquan, tai ji, or Tai Chi chuan, there are many different forms and styles of Tai Chi. 'Forms' refers to a routine of individually choreographed movements performed in a specific sequence and style (30).

There are several styles of Tai Chi (e.g. Chen, Yang, Sun, Wu and Hao) each of which have their own set of forms (or movements). Underlying all styles of Tai Chi are the fundamental principles of 'movement control' (soft, slow, continuous and precise rhythmic movements that flow with evenness and gentle resistance), 'body structure' (upright, supple posture maintained with a lowered centre of gravity and mindfulness for balance and weight transference) and 'internal components' (deep breathing and mental quietness to keep limbs and ligaments relaxed and enable awareness of the presence and movement of the body within its own space) (31-33).

Since the 1980s, Tai Chi has gained popularity in the West as a therapeutic and/or preventative physical activity (33). The practice of Tai Chi typically includes a warm-up followed by sets of movement, called forms or routines, which are performed in a certain way according to the style of Tai Chi being practiced. Forms may be described as 'short' or 'long', with the longer forms described as more physically and mentally demanding. Each style has its own established length of practice, and, like many exercise movement therapies, the proficiency and experience of an individual influences the speed, duration, and frequency of practice. For many, the aim is to maintain a daily practice, which can last anywhere from 5 minutes up to 2 hours (34). Tai Chi practice may also incorporate Qigong, an ancient Chinese form of exercise that aims to cultivate qi through breath, movement and meditation.

Tai Chi can be practiced at any time and in any location where there is sufficient space (35). It does not require specialist facilities, expensive equipment, or dedicated clothing and can be practiced regardless of age or level of fitness. In the Australian setting, Tai Chi is commonly taught and practiced in groups, both indoors and outdoors, usually under the instruction of a Tai Chi master. Individuals may also practice Tai Chi at home, while viewing or listening to professional Tai Chi videos or other multimedia. There is currently no regulatory method of certifying teachers or instructors in Tai Chi in Australia.

1.3 How the intervention might work

As a low-impact, moderately intense aerobic exercise involving complex motor control and deep diaphragmatic breathing, Tai Chi is thought to benefit the mind and body of practitioners by building strength from 'within' (10, 33, 35). This is achieved by combining isometric and isotonic exercise of deep stabiliser muscles and joints with relaxation, controlled breathing, and meditation (36). Isometric exercises involve static strength training, where specific muscle groups are contracted and held in one position, whereas isotonic exercises, involve repetitive lengthening or shortening of the muscle group in a way that causes joint movement. The potential physical benefits of Tai Chi are related to the regular practice of isotonic exercise, which enhances cardiopulmonary fitness through optimisation of oxygen utilisation, increasing exercise capacity and improving muscle strength, flexibility and endurance (35, 37, 38). In addition, the slow and deliberate weight-shifting feature of Tai Chi requires the muscles of the lower legs and feet to work considerably, as forms are completed with knees bent in a squat-like position. It is suggested that the intensity of Tai Chi (Yang style) does not exceed 50–55% of the individual's maximum oxygen intake (39-42) and changes in heart rate and blood pressure are similar to those for walking at a speed of 6 km/hr (43, 44).

The potential psychophysiological benefits of Tai Chi (such as improved attentiveness and sleep, and reduced fatigue, stress and anxiety), may be explained by the relaxation response theory that includes the homeostatic balance of sympathetic and parasympathetic nervous systems (45), reduced cellular inflammatory responses (46, 47), or through modulating or stimulating the connectivity of key brain regions involved in mood regulation and executive function (48-50). Further research is required to clarify and understand the mechanisms underlying the possible association between Tai Chi and other observed physiological responses such as reduced hypertension (51), improved immune function (52), and improved blood glucose control (53). Movement Qigong with meditation has similar proposed mechanisms of action, and therefore is expected to have the same benefits as Tai Chi (54).

1.4 Why it is important to do this review

In Australia, complementary therapies, including Tai Chi, 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 Tai Chi, to enable consumers, health care providers and policy makers to make informed decisions about care.

The 2015 review (4) identified 37 systematic reviews containing evidence from 117 unique RCTs involving 8852 participants across 16 clinical conditions and concluded that, compared with control, there is (a) very low quality evidence to suggest that Tai Chi may have some beneficial health effects in a limited number of conditions for a limited number of outcomes including the elderly (muscle strength), heart disease (quality of life), hypertension (SBP, DBP), and osteoarthritis (physical function), and (b) low to very low quality evidence that Tai Chi may have no effect on selected outcomes in people who are elderly (falls) and people with heart disease (HRV, exercise capacity).

Compared to active comparators, the 2015 review concluded that there is (a) very low quality evidence that Tai Chi may have beneficial effects relative to active comparators on selected outcomes in people with osteoarthritis (pain, physical function), and (b) very low quality evidence that suggests there may be no difference between Tai Chi and another active comparator in a limited number of conditions and for a limited number of outcomes including hypertension (SBP, DBP), osteoporosis (bone mineral density) and type 2 diabetes (HbA1c, FBG, total cholesterol).

The evidence for these findings is largely based on small, poor quality studies and was rated as very low for almost all outcomes (4). The magnitude and clinical significance of any potential health benefits were uncertain. For many outcomes, the health effects of Tai Chi were uncertain. A key limitation of the overview was the overall poor quality of information reported in the identified systematic reviews, and the implied poor quality of the RCTs they included.

In addition, four submissions were received, with two additional RCTs identified that were not identified by systematic reviews included in the overview (5). One RCT assessed Tai Chi in the elderly (community dwelling) and did not materially change the findings or interpretation of the evidence in the overview for this group. The other RCT assessed Tai Chi in people with low back pain, which was not included in the overview because there were no systematic reviews identified for this patient group. The interpretation of results for low back pain were therefore limited to this single study.

The rationale for conducting this review is to update and enhance the evidence and guidance used to inform the 2015 Overview of Tai Chi (4). That is, to identify whether any high-quality studies have been published since, or were not included in, the 2015 review, and address the evidence gaps noted. This is to ensure recommendations relating to the use of Tai Chi remain relevant and up to date.

2 Objectives

To conduct a systematic review of RCTs that evaluate the effectiveness of Tai Chi in individuals with a described injury, disease, medical condition, or preclinical condition. This will be supplemented with a systematic review of NRSIs for certain populations, settings or outcomes when a NRSI study design is more appropriate or feasible, in line with Cochrane recommendations (55).

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

3 Methods

Methods reported in this protocol are based on that described in the *Cochrane Handbook for Systematic Reviews of Interventions* (56) and relevant sections in the Joanna Briggs Institute Reviewer's manual (57). 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 collection, respectively. Where appropriate, RevMan (58) 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 RCTs that examine the effectiveness of Tai Chi compared to control or another intervention. As per Cochrane recommendations, NRSIs are only eligible for inclusion for certain populations, settings or outcomes that may be more appropriately or more feasibly evaluated using NRSI (55). The relevant populations, settings or outcomes will be determined via a blinded approach as specified in Sections 3.1.2 and 3.1.4.

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) (59). If the method of randomisation is not specifically stated, or not strictly random, then the study will be judged to be pseudorandomised. Pseudorandomised controlled trials 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.

For certain populations and/or outcomes (see Sections 3.1.2 and 3.1.4) NRSIs with design features as outlined in Table 1 are also eligible for inclusion. To be eligible for inclusion, the minimum design features of the NRSI 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 (60, 61).

NRSIs are included 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 (55). 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 (62), or the RCT evidence for a particular health condition and outcome is indirect and the question is better answered by available NRSI evidence (63).

Eligible NRSIs that are assessed to be at critical risk of bias for one or more domain (see Section 3.3.7) 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 1 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 (61, 64); Chapter 24 Including non-randomized studies on intervention effects (55); Cochrane Childhood Cancer (65);

Publication date

There are no limitations on publication date, however, studies published after the systematic review 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 pre-clinical 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 (66). For the purposes of this review, social determinants include factors such as 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, age, 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 may be included if separate data is available for the eligible population/s.

NRSIs will only be eligible for inclusion for certain populations, settings or outcomes as outlined in Section 3.1.1. (55) . These populations will be identified by the NTWC in parallel with the process for selecting critical and important outcomes (see Section 3.1.4). This will involve NTWC considering a list of the populations identified in included RCTs, while remaining blinded to the details and characteristics of the RCTs (e.g. risk of bias, outcome results). The NTWC will specify eligible populations from those listed, and will also nominate any other populations expected to be covered in the evidence review, for which RCTs may not be feasible, or where NRSIs may present the best available evidence due to the factors outlined in Section 3.1.1.

3.1.3 Types of interventions

Intervention

All styles and forms of Tai Chi are eligible for inclusion. That is, any activity in the name of Tai Chi instruction delivered to an individual or a group of individuals, or self-practiced.

There are no limits on intensity, duration of practice, or mode of delivery and studies will be included irrespective of whether the intervention is delivered by an instructor or through other media (e.g. instructional videos). Studies that include Tai Chi delivered in combination with Qigong will be included, however, studies in which the sole focus is Qigong will be excluded. Studies that include Tai Chi in combination with other forms of exercise will be excluded, unless the effect of Tai Chi alone can be discerned.

To allow for potential subgroup analysis (and to inform decision-making), studies will be stratified based on whether the participants receive instructor-led Tai Chi in a group or individual setting (see Section 3.3.15).

Comparators

There are no restrictions on the type of eligible comparators, noting that the analysis will stratify the evidence into two comparisons: (i) control (including no intervention, wait list or usual care); and (ii) other comparator (inclusive of exercise, education, and usual care if considered active).

Where usual care is poorly described or where usual care is described with Tai Chi as an adjunct (i.e. Tai Chi plus usual care vs usual care alone), it will be considered an inactive (control) 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 will be included if all arms of a study receive the same co-interventions (i.e. the effectiveness of Tai Chi is not confounded).

Restrictions: Studies comparing different styles, forms or components of Tai Chi with one another will be excluded.

3.1.4 Types of outcome measures

Outcome role

Outcomes will not be used as a criterion for including or excluding studies.

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 timepoints of interest

Any effectiveness outcome anticipated to demonstrate a treatment achieves its intended purpose is eligible for inclusion. 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 pre-specify outcomes. All pre-specified outcomes measured in each eligible 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' tables and corresponding evidence statements (see Section 3.3.17).

Outcome selection will occur after identification of eligible studies using a pre-specified 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 [http://www.comet-initiative.org/]).

Throughout the prioritisation exercise, the NTWC will remain blinded about the characteristics or results of included studies 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. At this time, the NTWC will also identify outcomes for which evidence from NRSIs will be considered, in line with the rationale provided in Section 3.1.1.

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 assessed as 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 RCTs and NRSIs and exclusions for publication types developed and published previously (67).

In developing the search strategy, we appraised and adapted the relevant search strategies provided in the 2015 review; with recent SRs 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 of English language databases. Non-English language databases will not be searched.

The search strategy will be adapted to suit the required syntax for the following electronic bibliographic databases:

- Embase (via Ovid)
- MEDLINE (via Ovid)
- Emcare (via Ovid) – coverage of all nursing specialty areas
- PsycINFO (via Ovid) – coverage of behavioural science and mental health
- AMED (via Ovid) – coverage of Allied and Complementary Medicine
- Cochrane Central Register of Controlled Trials (via Cochrane Library)
- PEDro – coverage of physiotherapy
- 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
- 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)

As the populations and/or outcomes for which NRSIs will be eligible will be specified by the NTWC after initial screening of RCTs (see Sections 3.1.2 and 3.1.4) the search strategy will be implemented in two phases. First, all eligible RCTs will be searched using the strategy outlined at Appendix A. Populations and/or outcomes eligible for inclusion of NRSIs will then be identified, and the search strategy at Appendix A will be augmented with population-specific search strings, to identify relevant NRSIs, with additional search terms approved by the NTWC prior to implementation.

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), however, any grey literature will be excluded.

3.3 Data collection and analysis

Included studies will be critically appraised, appropriate data extracted into data extraction tables, and the results analysed and summarised into appropriate categories or subgroups according to identified populations and 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. 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 relevance, 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 in the 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 inclusion, 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 the NTWC will be presented with excerpts from the publication relevant to the query while remaining blinded to the other identifying details such as study design, size, risk of bias, or results).

If a study 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. Eligible studies that are not available in English will be noted and managed as described in the below under *Studies published in languages other than English*.

Evidence provided through the Department's public call for evidence

Potentially relevant 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.

All of the submitted literature will be collated, tabulated, and cross referenced with the evidence identified via literature searches as described in Section 3.3.1. In-scope studies 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 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*').

3.3.2 Data collection process

For each included primary study, one reviewer will extract data using a standard pre-tested data extraction and coding form (see Appendix D). Pre-testing will involve all reviewers, who will each data extract the required information from the same three primary studies 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 as planned and any necessary revisions will be 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.

3.3.3 Requests for data

Eligible primary studies not published in English, ongoing trials and studies published as conference abstracts with incomplete results will be identified for inclusion. Study authors will be contacted through an open-ended request for further information. If no results are available, the study will be noted as '*Ongoing*' or, if information on study eligibility is lacking, the study will be recorded as '*Studies Awaiting Classification*' and will not be included in the evidence appraisal.

No attempts will be made to obtain or clarify data from published peer-reviewed studies.

3.3.4 Data items

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, frequency of

practice, program duration, co-interventions), outcomes (including measurement method, timing, or severity), and funding sources.

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 3.3.7).

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:

- Revised Cochrane Risk of Bias tool v2.0 (68, 69)
- ROBINS-I: a tool for assessing risk of bias in nonrandomised studies of interventions (70)

Randomised controlled trials

The risk of bias of RCTs will be assessed using the revised Cochrane Risk of Bias tool (68). 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 (70) using the ROBINS-I tool. Potential confounders and cointerventions for a 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 for risk of bias, 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*'.

3.3.7 Risk of bias assessment process

The risk of bias for each included study 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.

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

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.

3.3.8 Measures of effect

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 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 be reported as adjusted effect estimates (e.g. adjusted odds ratios (OR) from logistic regression or adjusted rate ratios from Poisson regression analyses). 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 pre-specify 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.

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.

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 results from more than one timepoint, results from a single timepoint will be selected for any given outcome, and only data from that timepoint will be presented in the analysis. The timepoint selected will be based on that determined to be 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 comparator (i.e. 'control' 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 (71).

3.3.11 Quantitative synthesis

Synthesis will only be undertaken for studies that compare Tai Chi with 'control'. Results data from studies comparing Tai Chi with 'other' interventions will be extracted and presented in data tables but will not be synthesised further, except where requested by the NTWC. These data will be presented as an 'evidence inventory' to provide a snapshot of the available evidence comparing Tai Chi with 'other' interventions.

The NTWC may request that data comparing Tai Chi with another active intervention be synthesised, where:

1) at least two studies compare the effect of Tai Chi 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.

Data from RCTs

Data synthesis from RCTs will be performed using RevMan 5.3 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 (72).

Effect estimates will not be combined across outcomes if analyses of covariance have been used to adjust for baseline measures, or for time-to-event data reported as hazard ratios.

Data from NRSIs

For those populations and/or outcomes for which NRSIs are included, data synthesis from NRSIs will be performed using RevMan 5.3 (where appropriate) and forest plots will be 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 (72).

Effect estimates will only be combined across outcomes if the included NRSIs are judged to be at low to moderate risk of bias (see Section 3.3.7) 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.

3.3.12 Non-quantitative synthesis

The evidence review will provide a structured, narrative summary of the results for each condition identified, along with risk of bias assessments, and other intervention characteristics, in tables structured by comparator ('control', or 'other' intervention), outcome domain, and study design (ordered and grouped by risk of bias, then 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.

The narrative summary will include a brief description of the condition and studies identified (including study design, size, and population demographics). This will be followed by a summary of results grouped by comparator and outcome domain. Result from each study will be reported, with the range and magnitude of observed effects noted. For studies where the 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. 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 (71).

To describe an overall effect across multiple studies for each outcome (for studies comparing Tai Chi with control only), we will describe the magnitude, range and distribution of observed effects across the studies using a simple vote count based on direction of effect (e.g. X/Y studies reported an effect favouring the intervention for the outcome Z). 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.

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 3.3.6, 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. '*Studies awaiting classification*') and discussed under '*Overall completeness and applicability of evidence*', noting that supplementary approaches for assessing bias due to missing studies (such as searching clinical trial registers, grey literature, or other reports) will not be performed.

If more than 10 RCTs are included for a particular PICO, funnel plots (of effect estimates against their standard errors) will be generated in RevMan 5.3 in order 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) (73). No additional statistical analysis for testing for small-study effects will be conducted.

3.3.14 Addressing risk of bias

All RCTs will be included in the review, regardless of judgements made regarding risk of bias. A description of the risk of bias of included studies in individual domains will be presented along with the effect estimate. 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 will not be included in the reporting of results, synthesis, and conclusion. 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*').

3.3.15 Subgroup analyses

We do not plan to undertake any subgroup analyses of subsets of participants within studies; however, if there is inconsistency between effect estimates, subgroup analysis will be used to explore possible sources of heterogeneity relating to the delivery of the intervention. Studies will be grouped according to intervention characteristics (i.e. intensity, duration, mode of delivery, or who delivers) and a standard test for heterogeneity across the subgroups will be reported.

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 from RCTs and NRSIs (where included) will be evaluated separately and only evidence comparing Tai Chi with 'control' will be presented.

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 across studies for each outcome reported for a comparison (74).
- *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 (75).
- *Imprecision.* Based on interpretation of the upper and lower confidence limits in relation to a minimal 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) (76).
- *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 (77).
- *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 (78).

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 3.3.17). 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, 79).

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).

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 from RCTs and NRSIs will be presented separately. Estimates of treatment effects for each outcome will be reported as absolute and relative risks (or SMD). 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 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 Tai Chi in [population] [is suggested to, may, results] in [little to no effect, reduce, increase, promote etc.] on [outcome] compared with [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 information:

- the methodology used to identify the evidence base (documented systematic literature search, inclusion and exclusion criteria described)
- the characteristics of included studies (data extraction and risk of bias forms)
- detailed results, presented by outcome, which will contain comprehensive 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 NHMRC NTWC members are lodged with the NHMRC and are available online.

This work is paid for 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. The NHMRC Clinical Trials Centre. Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies. Part A - Overview report for Tai Chi The University of Sydney 2014.
5. The NHMRC Clinical Trials Centre. Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies. Part B - Consideration of stakeholder submissions for Tai Chi The University of Sydney 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. Yang YL, Wang YH, Wang SR, Shi PS, Wang C. The effect of Tai Chi on cardiorespiratory fitness for coronary disease rehabilitation: A systematic review and meta-analysis. Frontiers in Physiology. 2018;8 (JAN) (no pagination)(1091).
8. Herrod PJJ, Doleman B, Blackwell JEM, O'Boyle F, Williams JP, Lund JN, et al. Exercise and other nonpharmacological strategies to reduce blood pressure in older adults: a systematic review and meta-analysis. Journal of the American Society of Hypertension. 2018;12(4):248-67.
9. Reychler G, Poncin W, Montigny S, Luts A, Caty G, Pieters T. Efficacy of yoga, tai chi and qi gong on the main symptoms of chronic obstructive pulmonary disease: A systematic review. Respiratory Medicine and Research. 2019;75:13-25.
10. Mudano AS, Tugwell P, Wells GA, Singh JA. Tai Chi for rheumatoid arthritis. Cochrane Database Syst Rev. 2019;9(9):CD004849.
11. Tai K, De Souza IMB, Marques AP. The effectiveness of Tai Chi Chuan in osteoarthritis of the knee: A systematic review. Annals of the Rheumatic Diseases. 2019;78 (Supplement 2):1060-1.
12. Stavale R, Jones F, Malavera A, Fregni F. Peripheral neuropathy. Neuromethods. 2018;138:399-443.
13. Cheng CA, Chiu YW, Wu D, Kuan YC, Chen SN, Tam KW. Effectiveness of Tai Chi on fibromyalgia patients: A meta-analysis of randomized controlled trials. Complementary Therapies in Medicine. 2019;46:1-8.
14. Wells RE, Beuthin J, Granetzke L. Complementary and Integrative Medicine for Episodic Migraine: an Update of Evidence from the Last 3 Years. Current Pain and Headache Reports. 2019;23 (2) (no pagination)(10).
15. Qin J, Zhang Y, Wu L, He Z, Huang J, Tao J, et al. Effect of Tai Chi alone or as additional therapy on low back pain: Systematic review and meta-analysis of randomized controlled trials. Medicine (Baltimore). 2019;98(37):e17099.
16. Charron S, McKay KA, Tremlett H. Physical activity and disability outcomes in multiple sclerosis: A systematic review (2011-2016). Multiple Sclerosis and Related Disorders. 2018;20:169-77.
17. Fidan O, Seyyar GK, Aras B, Colak E, Aras O. The effect of Tai Chi and Qigong on health-related quality of life in Parkinson's disease: a systematic review and meta-analysis of systematic reviews. Int J Rehabil Res. 2019;42(3):196-204.

18. Garcia-Rudolph A, Sanchez-Pinsach D, Salleras EO, Tormos JM. Subacute stroke physical rehabilitation evidence in activities of daily living outcomes: A systematic review of meta-analyses of randomized controlled trials. Medicine (Baltimore). 2019;98(8):e14501.
19. Zeng Y, Xie X, Cheng ASK. Qigong or Tai Chi in Cancer Care: an Updated Systematic Review and Meta-analysis. Current Oncology Reports. 2019;21 (6) (no pagination)(48).
20. Ni X, Chan RJ, Yates P, Hu W, Huang X, Lou Y. The effects of Tai Chi on quality of life of cancer survivors: a systematic review and meta-analysis. Supportive Care in Cancer. 2019;27(10):3701-16.
21. Kreutz C, Schmidt ME, Steindorf K. Effects of physical and mind-body exercise on sleep problems during and after breast cancer treatment: a systematic review and meta-analysis. Breast Cancer Research and Treatment. 2019;176(1).
22. Ford CG, Vowles KE, Smith BW, Kinney AY. Mindfulness and Meditative Movement Interventions for Men Living With Cancer: A Meta-analysis. Ann Behav Med. 2019;26:26.
23. Wayne PM, Lee MS, Novakowski J, Osypiuk K, Ligibel J, Carlson LE, et al. Tai Chi and Qigong for cancer-related symptoms and quality of life: a systematic review and meta-analysis. Journal of cancer survivorship : research and practice. 2018;12(2):256-67.
24. Zou L, Yeung A, Li C, Wei G-X, Chen KW, Kinser PA, et al. Effects of Meditative Movements on Major Depressive Disorder: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. 2018;7(8).
25. Sharma M, Haider T. Tai chi as an alternative and complimentary therapy for anxiety: a systematic review. Journal of evidence-based complementary & alternative medicine. 2015;20(2):143-53.
26. Pfeifer M, Dionyssiotis Y. Musculoskeletal Rehabilitation after Hip Fracture: A Review. Osteologie. 2019;28(3):183-91.
27. Lim KHL, Pysklywec A, Plante M, Demers L. The effectiveness of Tai Chi for short-term cognitive function improvement in the early stages of dementia in the elderly: a systematic literature review. Clinical interventions in aging. 2019;14:827-39.
28. Zou L, Loprinzi PD, Yeung AS, Zeng N, Huang T. The Beneficial Effects of Mind-Body Exercises for People With Mild Cognitive Impairment: a Systematic Review With Meta-analysis. Archives of Physical Medicine and Rehabilitation. 2019;100(8):1556-73.
29. Tai Chi Australia. What is Tai Chi? (Taijiquan) [2 January 2020]. Available from: http://www.taichiaustralia.com.au/TaiChi/what-is-tai-chi.
30. Tai Chi Australia. Tai Chi Forms Taught at Tai Chi Australia [January 2020]. Available from: http://www.taichiaustralia.com.au/TaiChi/tai-chi-forms.
31. Wolf SL, Coogler C, Xu T. Exploring the basis for Tai Chi Chuan as a therapeutic exercise approach. Arch Phys Med Rehabil. 1997;78(8):886-92.
32. Lam P. The Tai Chi Principles: Tai Chi for Health Institute; 2016 [02 January 2020]. Available from: https://taichiforhealthinstitute.org/essential-tai-chi-principles-2/.
33. Zhang HW, Zheng R, Xu C, Lin ZX, Qin Y, Cai J, et al. Tai Chi for improving recovery after stroke. Cochrane Database Syst Rev. 2018(3).
34. Lam P. Tai Chi for Health Institute. Frequently asked questions [Available from: https://taichiforhealthinstitute.org/faqs/#practice.
35. Ng SM, Wang CW, Ho RT, Ziea TC, He J, Wong VC, et al. Tai chi exercise for patients with heart disease: a systematic review of controlled clinical trials. Altern Ther Health Med. 2012;18(3):16-22.
36. Larkey L, Jahnke R, Etnier J, Gonzalez J. Meditative movement as a category of exercise: implications for research. J Phys Act Health. 2009;6(2):230-8.
37. Lam P. Research into the effects of Tai Chi 2012. Available from: https://www.taichiaustralia.com/articles/2012-12-22_Sub_to_NHMRC.html.

38. Sam-Kit Tin T, Daniel Weng C-H, Vigário PdS, Ferreira AdS. Effects of A Short-term Cardio Tai Chi Program on Cardiorespiratory Fitness and Hemodynamic Parameters in Sedentary Adults: A Pilot Study. Journal of Acupuncture and Meridian Studies. 2020;13(1):12-8.

39. Zhuo D, Shephard RJ, Plyley MJ, Davis GM. Cardiorespiratory and metabolic responses during Tai Chi Chuan exercise. Can J Appl Sport Sci. 1984;9(1):7-10.

40. Lan C, Chou SW, Chen SY, Lai JS, Wong MK. The aerobic capacity and ventilatory efficiency during exercise in Qigong and Tai Chi Chuan practitioners. The American journal of Chinese medicine. 2004;32(1):141-50.

41. Lan C, Chen SY, Lai JS. The exercise intensity of Tai Chi Chuan. Medicine and sport science. 2008;52:12-9.

42. Taylor-Piliae RE. The effectiveness of Tai Chi exercise in improving aerobic capacity: an updated meta-analysis. Medicine and sport science. 2008;52:40-53.

43. Jin P. Efficacy of Tai Chi, brisk walking, meditation, and reading in reducing mental and emotional stress. J Psychosom Res. 1992;36(4):361-70.

44. Yeh GY, Wang C, Wayne PM, Phillips RS. The Effect of Tai Chi Exercise on Blood Pressure: A Systematic Review. Preventive Cardiology. 2008;11(2):82-9.

45. Yeung A, Chan JSM, Cheung JC, Zou L. Qigong and Tai-Chi for Mood Regulation. Focus (Am Psychiatr Publ). 2018;16(1):40-7.

46. Irwin MR, Olmstead R. Mitigating cellular inflammation in older adults: a randomized controlled trial of Tai Chi Chih. Am J Geriatr Psychiatry. 2012;20(9):764-72.

47. Irwin MR, Olmstead R, Breen EC, Witarama T, Carrillo C, Sadeghi N, et al. Tai chi, cellular inflammation, and transcriptome dynamics in breast cancer survivors with insomnia: a randomized controlled trial. J Natl Cancer Inst Monogr. 2014;2014(50):295-301.

48. Kong J, Wilson G, Park J, Pereira K, Walpole C, Yeung A. Treating Depression With Tai Chi: State of the Art and Future Perspectives. Front Psychiatry. 2019;10:237.

49. Cui L, Yin H, Lyu S, Shen Q, Wang Y, Li X, et al. Tai Chi Chuan vs General Aerobic Exercise in Brain Plasticity: A Multimodal MRI Study. Sci Rep. 2019;9(1):17264.

50. Pan Z, Su X, Fang Q, Hou L, Lee Y, Chen CC, et al. The Effects of Tai Chi Intervention on Healthy Elderly by Means of Neuroimaging and EEG: A Systematic Review. Front Aging Neurosci. 2018;10:110.

51. Campo RA, Light KC, O'Connor K, Nakamura Y, Lipschitz D, LaStayo PC, et al. Blood pressure, salivary cortisol, and inflammatory cytokine outcomes in senior female cancer survivors enrolled in a tai chi chih randomized controlled trial. J Cancer Surviv. 2015;9(1):115-25.

52. Irwin MR, Olmstead R, Oxman MN. Augmenting immune responses to varicella zoster virus in older adults: a randomized, controlled trial of Tai Chi. J Am Geriatr Soc. 2007;55(4):511-7.

53. Chao M, Wang C, Dong X, Ding M. The Effects of Tai Chi on Type 2 Diabetes Mellitus: A Meta-Analysis. J Diabetes Res. 2018;2018:7350567.

54. Jahnke R, Larkey L, Rogers C, Etnier J, Lin F. A comprehensive review of health benefits of qigong and tai chi. Am J Health Promot. 2010;24(6):e1-e25.

55. 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.

56. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019): Cochrane; 2019. Available from: www.training.cochrane.org/handbook.

57. 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.

58. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration; 2014.

59. 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.

60. 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.

61. 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.

62. 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.

63. 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.

64. 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.

65. Cochrane Childhood Cancer. 2020 [Available from: https://childhoodcancer.cochrane.org/non-randomised-controlled-study-nrs-designs.

66. Australian Institute of Health and Welfare. Australia's health 2016. Canberra; 2016.

67. National Blood Authority. Patient Blood Management Guidelines: Module 6 Neonatal and Paediatrics. Technical report - Volume 1. Canberra, Australia: National Blood Authority; 2016. Available from: https://www.blood.gov.au/pbm-module-6.

68. 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.

69. 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.

70. 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.

71. 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).

72. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.

73. 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.

74. 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.

75. 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.

76. 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.

77. 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.

78. 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.

79. 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.

Appendix A - Literature search strategy

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

1. 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.

2. 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.

3. case report/
4. (editorial or letter or comment or historical article).pt.
5. (animals/ or nonhuman/) not humans/
6. 3 or 4 or 5

Concept: Tai Chi

7. exp Tai Ji/
8. (tai ji or tai-ji or taiji or tai?ji).ti,ab.
9. (Tai Chi or t'ai chi or thai chi or t?ai chi).ti,ab.
10. (tai ji quan or taijiquan).ti,ab.
11. (tai zi zhang or taijizhang).ti,ab.
12. (Tai Chi chuan or taichichuan).ti,ab.
13. Ai Chi.ti,ab.
14. or/7-13

Concept: evidence hierarchy for screening

15. (14 AND 1) NOT 6
16. (14 AND 2) NOT 6*

*Population-specific search terms will be added to this search once populations and/or outcomes eligible for NRSIs are specified by NTWC (see Section 3.2.1).

The above strategy will be adapted to suit EBSCO (CINAHL, AMED), 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

CINHAL 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.

At abstract/title screening items 1 through 8 will be considered and applied.

At full text review, all items will be considered and applied as appropriate (these studies will be listed in the technical report with reasons for exclusion).

1. duplicate citation
2. nonhuman study
3. intervention out of scope (not Tai Chi or component thereof)
4. comparator out of scope (compares different styles, forms or components of Tai Chi)
5. population out of scope (healthy participants seeking general wellness)
6. outcome out of scope (patient experience, 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 [a] (specify)
    a. non-systematic review, Guideline, HTA assessment
    b. SR of RCTS or NRSIs
    c. case series or other
9. duplicate citation submitted to the Department (RCT / NRSI 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 (Study has been updated or 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

a. Screening of articles not published in English will be conducted as described in the Section "*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

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 randomization 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 were 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 pre-specified 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 of the RCT		

Abbreviations:

Source: Adapted from Chapter 8 Cochrane Handbook for Systematic Reviews of Interventions (68).

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 co-interventions 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.
Bias in measurement of outcomes	Low risk Moderate risk Serious risk Critical risk	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;

	No information	• 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 (70)

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

Study ID	**Author date**				
Study design	RCT/NRSI		Features:		
Affiliation/source of funds	Source of funding and conflicts of interest not declared				
Enrolment period	Aug 2014 to Mar 2016				
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, frequency, program duration)				
Intervention	n=				
Comparator #1 (control)	n=				
Comparator #2 (other)	n=				
Co-interventions					
add rows as needed	n=				
Is instructor certified? (select from list)	Yes	No	Not specified		
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		
Missing data	e.g. imputations, loss to follow-up				

INTERNAL VALIDITY	
Overall risk of bias (select from list)	Some concerns for one or more domains, but no high risk of bias
Summary (descriptive)	

EXTERNAL VALIDITY	
	The evidence is directly generalisable to the Australian population

Study ID	Author date
Generalisability (relevance of the study population to the Australian population)	e.g. The study was conducted in elderly patients, mean age is comparable to that in the Australian population
Applicability (relevance of the evidence to the Australian health care system)	The evidence is directly applicable to the Australian health care context with few caveats
	e.g. The study was conducted in Italy and is likely to be relevant to the Australian health care context with the exception that ….? This is not expected to influence the outcomes of the study

Eligibility for this review				
Criterion	(select from dropdown)			
Population meets eligibility criteria	Uncertain	Seek guidance from NTWC		
Service or service component as practiced in Australia	Uncertain	Seek guidance from NTWC		
At least one relevant outcome measure	Uncertain	Seek guidance from NTWC		

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

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[a]	Direction of effect
			[outcome #1]	< 3 months from randomisation	e.g. VAS scale 1-100	higher score means more pain				RR 1.00 [0.68, 1.48]		*No difference*
			[outcome #2]	(> 3 months but less than 1 year from randomisation)								*Favours intervention*
			[outcome #3]	(> 1 yr from randomisation)								*Not reported*
			[outcome #4]									
			[outcome #5]									
			[outcome #6]									
			[outcome #7]									

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

Notes:

a. Data as reported by the study authors.