

Systematic review of evidence on the clinical effectiveness of kinesiology

Protocol prepared by
Cochrane Australia

21 March 2024



Contents

Authors and contributors to the protocol	4
Declarations of interest	4
Funding	4
1. Background	5
1.1 Description of the intervention	5
1.2 How kinesiology might work	6
1.3 Description of conditions for which kinesiology is used	7
1.4 Why it is important to do this review	7
2. Objectives	7
Primary objectives	8
Secondary objectives	9
3. Methods	9
3.1 Criteria for considering studies for this review	10
3.1.1 Types of studies	10
3.1.2 Types of participants	12
3.1.3 Types of interventions	13
3.1.4 Types of outcomes	16
3.2 Search methods for identification of studies	18
3.2.1 Electronic searches	18
3.2.2 Selection of search terms	18
3.2.3 Searching other resources	18
3.3 Data collection and analysis	19
3.3.1 Selection of studies	19
3.3.2 Data extraction and management	20
3.3.3 Assessment of risk of bias of included studies	22
3.3.4 Measures of treatment effect	23
3.3.5 Unit of analysis issues	24
3.3.6 Dealing with missing data	25
3.3.7 Assessment of heterogeneity	25
3.3.8 Assessment of biases due to missing results	25
3.3.9 Data synthesis	26
3.3.10 Summary of findings tables and assessment of certainty of the body of evidence	27

4. References	30
Appendix 1: Database and register search strategies	34
Cochrane Central Register of Controlled Trials	34
MEDLINE via Ovid	34
Embase via Ovid	34
AMED via Ovid	34
Emcare via Ovid	34
CINAHL	35
Europe PMC	35
ClinicalTrials.gov and WHO ICTRP	35
Appendix 2: Study design features checklist	36
Appendix 3: TIDIER domains and example of application in kinesiology SR	38

In April 2022 Cochrane Australia was contracted by the National Health and Medical Research Council (NHMRC) to design and undertake the systematic review described in this protocol. This systematic review is one of several independent contracted evidence evaluations being undertaken to update the evidence underpinning the 2015 *Review of the Australian Government Rebate on Natural Therapies for Private Health Insurance* (2015 Review) by the Department of Health and Aging (Department). The design and conduct of the review will be done in collaboration with the Office of NHMRC (ONHMRC), the NHMRC's Natural Therapies Working Committee (the Committee) and Department's Natural Therapies Review Expert Advisory Panel (NTREAP).

Authors and contributors to the protocol

Max Murano ¹ max.murano@monash.edu	Conducted background research on kinesiology (including characterising the interventions, mechanism of action and scope of practice, conditions treated and outcomes), drafted the related sections of the background and intervention eligibility criteria. Will co-lead review, develop data collection tools and contribute to acquisition of data.
Steve McDonald ¹ steve.mcdonald@monash.edu	Designed and developed the search strategy and wrote the search methods. Will conduct and report the search, and contribute to acquisition of data from studies.
Joanne McKenzie ² joanne.mckenzie@monash.edu	Designed and wrote the analysis plan and method for reporting treatment effects. Wrote the section on Assessment of biases due to missing results. Provided statistical advice on risk of bias assessment and interpretation. Planned contributions to the analysis and interpretation.
Sue Brennan ^{1*} sue.brennan@monash.edu	Senior Evidence Officer responsible for leading the review. Led design of the review and co-led writing of the protocol, with contributions from other authors as described.

¹ Cochrane Australia, Evidence Synthesis Qualitative & Implementation Methods Unit, School of Public Health and Preventive Medicine, Monash University, St Kilda Rd, Melbourne, Australia

² Methods in Evidence Synthesis Unit, School of Public Health and Preventive Medicine, Monash University
* contact author

Declarations of interest

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

Funding

This review is funded by the Australian Government Department of Health and Aging via the NHMRC, under Contract Reference ID: 2021-22C021.

© The Author(s). 2020 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

1. Background

In 2015, the Australian Government conducted a Review of the Australian Government Rebate on Natural Therapies for Private Health Insurance (2015 Review). Underpinned by systematic reviews of evidence for each natural therapy, one of the findings from the 2015 Review was that there was insufficient evidence to reach any conclusion that specialised (energy) kinesiology was effective. 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. This evidence evaluation of kinesiology is one of a suite of independent contracted systematic reviews that will inform the Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies 2019-20 (2019-20 Review) [1].

The term ‘kinesiology’ is used in multiple ways that fall outside the scope of this review, referring to the science of human movement (i.e. the anatomical, physiological, biomechanical and psychological elements of movement), a specific therapeutic tool used in physiotherapy practice (kinesiology taping), and the ‘applied kinesiology’ techniques, developed by Dr George Goodheart DC in the 1960s to augment the standard methods of diagnosis used by Chiropractors, for which practitioners in Australia must hold a university qualification in a relevant healthcare discipline.

This review focuses on specialised kinesiology. Specialized kinesiology originated from ‘Touch for Health’ kinesiology (T4H) which emerged in the 1970's as a simpler form of applied kinesiology that was intended for self-use outside the chiropractic and other health professions [2, 3]. Since then, a number of other systems of specialised, or energy, kinesiology have been developed based on a diverse range of energetic, counselling, education and healing concepts. Variants of specialised, or energy, kinesiology systems include Professional Kinesiology Practice (PKP), Three in One, Kinergetics, and integrative kinesiology [4, 5]. For the purposes of this review, we use the term 'kinesiology' to refer to any specialised, or energy, kinesiology system or method.

In Australia, the main source of information about the rates of consultation with complementary medicine practitioners is a cross-sectional survey conducted as part of the Practitioner Research and Collaborative Initiative (PRACI) [6]. The 2017 PRACI survey of Australian adults found that about a third of all respondents (36%; 726/2025 respondents) had consulted at least one complementary therapist in the last 12 months. Respondents were not asked whether they had consulted a kinesiologist.

1.1 Description of the intervention

The Australian Kinesiology association describes kinesiology as a non-invasive holistic therapy that uses manual muscle testing to assess imbalances expressed in the body (i.e. “anatomical, physiological and psychological stressors”) [7]. The muscle testing results inform the selection of individualised healing components from a wide variety of modalities by identifying “the elements which inhibit the body’s natural internal energies” and accessing “the life enhancing potential within the individual” [4, 7]¹. From a description provided by Jensen 2012, “During a muscle test, a practitioner applies a force to one muscle or group of muscles, with a particular intent in mind. The muscle is then labelled “weak” or “strong” based on its ability to resist this force.” [8]. Although specialised kinesiology is a derivative of Applied Kinesiology, the muscle testing

¹ Definition approved by the Australian Kinesiology Association 1999, amended in 2006. <https://aka.asn.au/what-is-kinesiology/>

techniques used in latter are highly systematised, requiring practitioners to follow a consistent protocol in order to meet requirements of the International College of Applied Kinesiology. The extent to which specialised kinesiology follows systematised protocols is unclear and such derivatives have been described as lacking ‘one or more of the essential attributes’ of Applied Kinesiology [9].

Drawing on the testing responses, which is said to “identify the bottom line cause of any imbalance”, the kinesiologist may select individualised healing components from a wide variety of tools or modalities to “facilitate the natural healing process” [7]. Some examples of kinesiology 'tools of the trade' are: Bach flower remedies; nutritional and/or dietary supplements and advice; acupuncture; aromatherapy; posture, alignment and myofascial interventions; and more psychotherapeutic interventions, such as Emotional Freedom Techniques (EFT), Eye Movement Desensitisation and Reprocessing (EMDR), goal setting and visualisations (as identified by the Australian Kinesiology Association [10]). These techniques are used to “correct the underlying energies in the body that allows the body, itself, to heal” rather than to “treat any specific disorder, disease or symptom” [7].

Mode of administration and dose

During a kinesiology session, the practitioner uses manual muscle testing to identify blockages or imbalances. Healing modalities may then be selected and deployed in the session and/or discussed with the client for home use. Some kinesiologists also offer online consultations, and describe using ‘surrogate’ muscle testing, where the practitioner tests themselves on behalf of the client to “identify any stressors, blockages and imbalances” [11]. Kinesiology sessions can last from 30 to 90 minutes. Professional associations indicate that some clients only require a single session, whereas others require multiple sessions [12].

Practitioners of kinesiology and regulation

In Australia, kinesiology is usually delivered by practitioners who have attended accredited kinesiology training, ranging from entry-level courses through to Certificate IV (~450 hours), Diploma (1200 hours) and Advanced Diploma (2150 hours) levels. The Australian Institute of Kinesiologists Ltd (AIK) and the Australian Kinesiology Association (AKA) accredit Registered Training Organisations (RTOs), Modality Owners (MOs) or other practitioners to deliver workshop units and training courses in core kinesiology and/or tools of the trade and/or associated health courses such as nutrition, anatomy, physiology, mind body medicine, communication and counselling. Both organisations offer varying levels of membership to kinesiology practitioners based on participation in requisite accredited training and courses, coordinate continuing professional development programs, and have developed Codes of Ethics and Codes of Practice for their members [13, 14].

The practice and teaching of kinesiology is not regulated by the Australian Health Practitioner Regulation National Law, which means there is no requirement for professional registration of practitioners of kinesiology [15]. This is an important distinction between the forms of kinesiology eligible for this review, and applied kinesiology for which practitioners in Australia must hold a university qualification in a relevant healthcare discipline that is eligible for registration under the Australian Health Practitioner Regulation National Law (e.g. chiropractic, osteopathy, physiotherapy or medicine) and have completed the study requirements to become a member of the International Board of Applied Kinesiology.

1.2 How kinesiology might work

Kinesiology professional associations and practitioners of a range of kinesiology systems describe engaging the electrical systems of the body during muscle testing as a form of biofeedback. It is thought that muscle testing identifies blockages or disruptions to energy flow throughout the body, allowing the practitioner and client to identify issues or challenges across mind, body or spirit. The muscle testing also indicates the tools needed to correct these imbalances, and so address or remediate the issues. The various tools and techniques described are said to target the body's structural, nutritional, electrical, chemical, mental, emotional or other energy systems. Each of these tools and techniques draws from a variety of philosophic and therapeutic approaches with their own proposed mechanisms of action [4, 5, 7, 16].

1.3 Description of conditions for which kinesiology is used

The Australian Kinesiology association state that “kinesiologists do not diagnose or treat any specific disorder, disease or symptom” [7]. However, kinesiology is sought by people with many different health conditions. Respondents to an Australian survey of practitioners with a qualification in kinesiology reported that they ‘often’ treated fatigue (19/19 respondents, 100%), digestive disorders / irritable bowel syndrome (94% / 72%), mental illness (84%), menstrual disorders (72%), insomnia/sleeping disorders (68%), thyroid complaints (61%), chronic pain (50%), recurrent infections (50%), and headache (42%) [17]. A majority of respondents also reported that they sometimes or often treated arthritis, allergic rhinitis (hayfever), attention deficit hyperactivity disorder (ADHD)/autism, eczema and psoriasis, and asthma.

Australian professional organisations suggest kinesiology as a treatment for stress and mental distress, pain, learning difficulties, cardiovascular issues, digestive issues, skin disorders, food allergies or intolerances and nutritional imbalances among other health conditions [7, 18].

1.4 Why it is important to do this review

This systematic review will inform the Australian Government’s Natural Therapies Review 2019-20, which is evaluating evidence of the clinical effectiveness of 16 therapies (including kinesiology). The conclusion from the evidence evaluation conducted on kinesiology for the *2015 Review* was that it “is insufficient evidence from [systematic reviews] within this field to reach any conclusion regarding the effectiveness, safety, quality or cost-effectiveness of kinesiology” [19]. The evidence evaluation used overview methods, identifying a single systematic review published since April 2008 [4]. None the primary studies included in this systematic review were RCTs, and therefore no primary evidence was considered in the overview. In contrast to the 2015 kinesiology evidence evaluation, this review will examine evidence from eligible primary studies published from database inception until the date of the last search for this systematic review. This approach will identify both new primary studies and those in areas that had not been addressed in a systematic review eligible for inclusion in the overviews conducted for the *2015 Review*.

2. Objectives

The overall objective of this systematic review is to examine the evidence for the clinical effectiveness of kinesiology in preventing and/or treating injury, disease, medical conditions or preclinical conditions [1]. The review will focus on outcomes (and underlying conditions) for which kinesiology is commonly sought or prescribed in Australia, and which are relevant to the 2019-20 Review of the Private Health Insurance rebate.

The specific objectives of the review follow (framed as questions). Examples of potentially relevant outcome domains and conditions are included to illustrate the breadth of questions to be addressed in the synthesis. These questions will be refined through a staged prioritisation process

(Methods, Figure 1) to align with priorities for the 2019-20 Review, ensure a consistent approach across the evidence evaluations of natural therapies (where appropriate), and make best use of available evidence.

Primary objectives

1. What is the effect of *kinesiology* compared to *no kinesiology (inactive controls)* (see section 3.1.3 Comparisons) among people with any condition, pre-condition, injury or risk factor on outcomes for which kinesiology is indicated? (for example, what is the effect on pain, anxiety or health-related quality of life irrespective of the underlying condition?)

Secondary objectives

2. For each specified condition, pre-condition, injury or at-risk population, what is the effect of *kinesiology* compared to *no kinesiology* on outcomes for which kinesiology is indicated? (for example, what is the effect on pain among people with chronic pain, dysmenorrhea, or episodic headache?)
3. What are the effects of *kinesiology* compared to *evidence-based 'gold standard' treatments*? (see 3.1.3 Types of interventions - Comparisons)
4. What evidence exists examining the effects of *kinesiology* compared to *other active comparators*? (i.e. not a 'gold standard')

3. Methods

Methods reported in this protocol are based on the *Cochrane Handbook for Systematic Reviews of Interventions* [20]. The GRADE approach will be used to summarise and assess the certainty of evidence arising from this review (see Section 3.3.9 for details). GRADE methods are widely used in systematic reviews and guideline development to ensure a systematic, transparent and common approach to interpreting results [21]. The protocol is reported in accordance with the PRISMA-P statement [22, 23] with consideration given to the extensively updated guidance for reporting methods for systematic review in the PRISMA 2020 statement [24, 25]. The review has been prospectively registered on the International prospective register of systematic reviews (PROSPERO ID CRD42024528900). Any changes to the protocol (including methods not used) will be documented when reporting the methods of the completed review.

The methods for this review are designed to accommodate the breadth of evidence about the effects of kinesiology relevant to the 2019-20 Review, and ensure a consistent approach with the other evidence evaluations of natural therapies (where appropriate). To achieve this, we will follow the staged approach summarised in Figure 1 and elaborated in subsequent sections. We begin with an initial analytic framework (step 1) that will be refined through a prioritisation process. To facilitate this process, we will screen studies against the review eligibility criteria and compile an aggregate list of outcomes for each population group, derived from the included studies and organised by the initial framework (step 2). No identifying information will be included (i.e. no study-level information, results, references, number of studies etc). The NTCW and NTCREAP will review the list in order to prioritise outcomes and advise on the final framework for the synthesis (step 3), which will be finalised (step 4) prior to proceeding with the review (step 5).

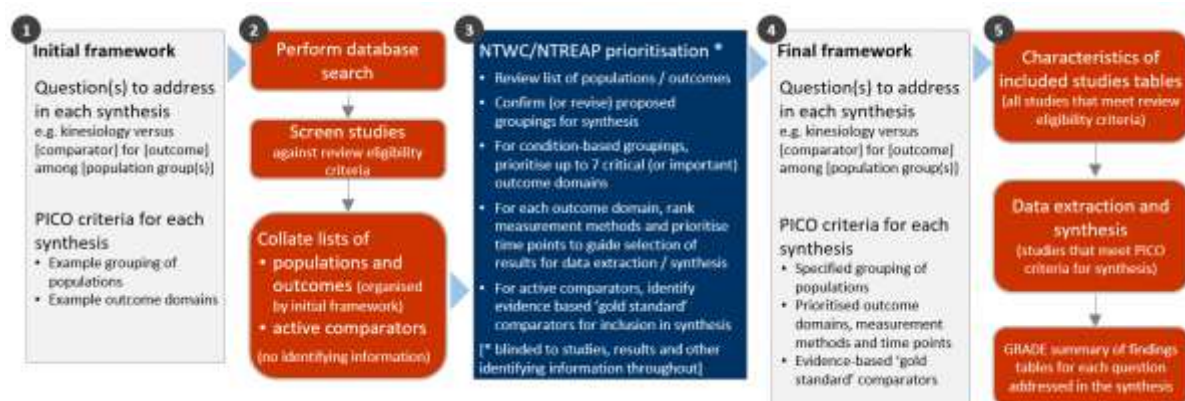


Figure 1. Staged approach for developing the analytic framework for this review

Figure 2 shows the initial analytic framework for the review. Example populations and outcome domains are included to convey the breadth of the review, and illustrate possible population and outcome groups for synthesis. These are indicative and not intended to be exhaustive. The framework was informed by research on the outcomes (and underlying conditions) for which kinesiology is commonly sought or prescribed, a scoping search of studies evaluating kinesiology, the wider literature on kinesiology, and consideration of frameworks for classifying disease and outcomes [26, 27]. Details for each PICO (population, intervention, comparator, outcomes) element follow (Section 3.1).

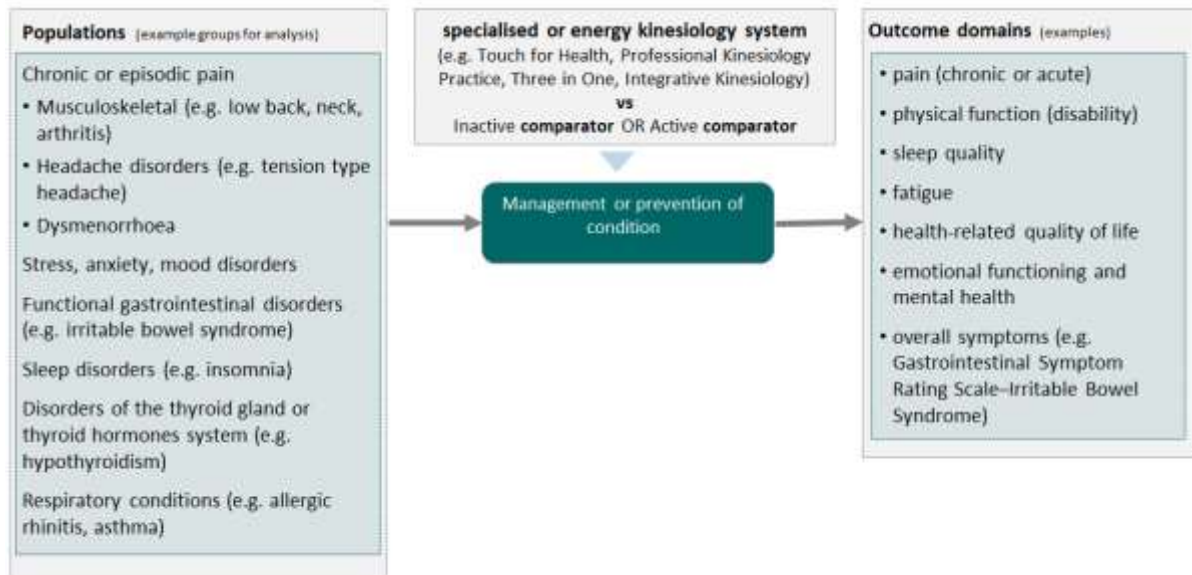


Figure 2. Initial analytic framework for the review and synthesis

3.1 Criteria for considering studies for this review

3.1.1 Types of studies

Randomised controlled trials (RCTs) and non-randomised studies of interventions (NRSIs) are eligible for inclusion in the review.

Randomised controlled trials including individually and cluster randomised, and cross-over trials are eligible for inclusion in the review.

Controlled trials in which the allocation sequence does not include a truly random element, was predictable, or was not adequately concealed from investigators, will be considered alongside RCTs when assessing risk of bias (i.e. using the Cochrane ROB 2 tool) and in the synthesis as long as there was an attempt to have some kind of ‘randomisation’ to groups. These studies are sometimes referred to as ‘quasi-random’. Examples include studies using methods for sequence generation based on alternation, dates (of birth or admission) and patient record numbers [45]. This is to ensure a consistent approach with other natural therapies systematic reviews.

Non-randomised studies of interventions (NRSIs) with specific design features are eligible for inclusion in the review. In line with current Cochrane guidance, only NRSIs with design features that are suitable for estimating a causal effect will be eligible. While study design labels will be used as an aid to communicating about eligible designs and for use in the review (Table 1), eligibility decisions will be based on assessment of the specific design features of each study rather than the label used by the study authors (see checklist Appendix 2) [28, 29].

Eligible designs are those in which the following features are present.

- The intervention may be allocated to individuals or clusters. We anticipate that kinesiology (or the control) will be allocated to individuals in most studies, although clustering is likely in these studies given the way in which kinesiology lessons are delivered (i.e. the same teacher may deliver the intervention to multiple participants) [30].
- Treatment groups may be formed by some action of the researchers or in the course of usual treatment decisions (including healthcare decision makers, practitioners or participants/patients/peoples' choices).
- Studies must include a contemporaneous control.
- There must be an attempt to control for confounding (either by using methods that control in principle for confounding or that control for observed covariates)
- The design must be suitable for estimating a causal effect.

Table 1. Study design labels to be used in the review with indicative features

Study design label	Thumbnail sketch of indicative features (extracts from [29, 31])
Nonrandomised controlled trial (NRCT)	Study in which allocation to intervention and comparator is not random or quasi-random (including where based on the practitioners' or participants' preferences) and is applied by research personnel. Participants are followed up from the start of intervention up to a later time for ascertainment of outcomes of interest [29, 31].
Prospective cohort study (PCS)	Study in which subjects are identified prospectively and classified as having received the intervention or comparator of interest on the basis of the prospectively collected information. Participants are followed up from the start of intervention up to a later time for ascertainment of outcomes of interest [29, 31].
Retrospective cohort study (RCS)	Study in which subjects are identified from historic records and classified as having received the intervention or comparator of interest on the basis of the historic information. Participants are followed up from the start of intervention up to a later time for ascertainment of outcomes of interest [29, 31].
Controlled Interrupted time series (CITS)	Studies that "collect longitudinal data measured at an aggregate level (across participants within one or more units), with several measurement times before implementation of the intervention, and several measurement times after implementation of the intervention." [31] Must have a contemporaneous control in which the intervention is not implemented to be eligible.
Controlled before-and after-study (CBA)	"Studies in which: (i) units are non-randomly allocated to a group that receives an intervention or to an alternative group that receives nothing or a comparator intervention; and (ii) at least one measurement of the outcome variable is made in both groups before and after implementation of the intervention..." [31]

Exclusions. Historical case control (no contemporaneous control), uncontrolled before-after studies (no contemporaneous control), and cross-sectional studies (not suitable for establishing a causal effect because both the outcome and exposure to the intervention are measured at the same time). Case-control studies are also ineligible because they are only suited to examining rare outcomes (which are unlikely to be among the core outcomes for the conditions for which kinesiology is used) and harms (which are outside the scope of this review).

Date and language restrictions.

There are no restrictions on publication date.

Potentially eligible studies published in languages other than English will not be included in the review, but will be listed according to whether they are likely to be eligible or whether eligibility cannot be determined (see 3.3.1 Selection of studies). The impact of excluding these studies will be considered in the assessment of bias due to missing results (see 3.3.8 Assessment of reporting bias, 3.3.9 Summary of findings tables and assessment of the certainty of the body of evidence).

3.1.2 Types of participants

Studies involving participants with any disease, medical condition, injury, or preclinical condition are eligible for the review. This includes healthy participants with clearly-identified risk factors (e.g. biomedical, health behaviours, or other). There are no restrictions on age or other demographic factors.

For trials in which kinesiology is used for primary or secondary prevention, participants must have a clearly-identified factor that puts them at heightened risk of the condition or injury that the intervention is intended to prevent compared to the population at large. Where possible, decisions about whether a population is at risk will be informed by evidence from a systematic review of risk factors.

We will operationalise this as follows:

- The risk factor(s) for the condition that kinesiology is used to prevent must be part of the eligibility criteria for the trial or reported in the baseline data (e.g. work that involves demanding posture or repetitive movement in a trial aiming to prevent workplace-related musculoskeletal conditions), and
- There must be a direct link between the risk factor and the trial outcomes (i.e. an outcome that demonstrates progression to a diagnosable condition or pre-condition; musculoskeletal pain or injury in a trial that aims to prevent injury)

We expect that studies will include participants that fall within broad population groups, such as those shown in **Figure 2**. These are indicative groups, included to illustrate the breadth of populations eligible for the review and possible groupings for synthesis. Decisions about which groups to include in the final analytic framework will be made through the prioritisation process (**Figure 1**). This process may lead to changes and additions to the population groups (i.e. broader, narrower or new groups).

Exclusions. Healthy populations seeking health improvement. This includes healthy participants using kinesiology to enhance general well-being. For borderline decisions, the NTWC will be given the study PICO, aims and any information about potential risk factors reported by the trialists (without results or information that identifies the study) and asked to adjudicate.

Studies that include both healthy participants and participants eligible for the review, will be included if separate data are available or a majority of participants meet the review eligibility criteria as per guidance in the Cochrane handbook [32]. For the latter, we will consider implications for the applicability of study findings in the GRADE assessment.

While studies involving any population will be included in the review (except for the specific exclusions above), if the number of eligible studies for synthesis is unmanageable, the synthesis may be limited to populations (conditions) most relevant to the use of kinesiology in Australia. Such decisions will be made through a similar process to that outlined in Figure 1, by the NTWC without knowledge of study results or other identifying information. The decisions will be guided by data about practice in the Australian context (e.g. practitioner or patient surveys that report reasons for use in Australia). Studies excluded from the synthesis to limit scope will be included in an evidence inventory (Objective 4).

3.1.3 Types of interventions

Because of the potential challenge of distinguishing components of specialised or energy kinesiology systems from related modalities such as applied kinesiology, and the likelihood of identifying studies in which the defining techniques and principles of specialised or energy kinesiology systems are incompletely reported, studies will be included if the therapy is described as a specialised or energy kinesiology system, or any of the named variants of this system (Touch for Health, Professional Kinesiology Practice (PKP), Three in One, Integrative Kinesiology).

For the purpose of this review, eligible interventions are specialised or energy kinesiology as per the description provided by the NHMRC [33]:

A system that involves the use of manual muscle testing techniques

- “... to elicit a yes/no response in muscles via a strong/ weak bio-feedback mechanism ... to reveal imbalances within the body; for example, sources of pain, allergies, and digestive and mental health disturbances”, and
- “facilitate a person’s natural healing process” by bringing “the root cause of any ‘imbalance’ to a person’s conscious attention” and identifying tools needed to correct any imbalances (extracts from [33]).

The latter is an essential component that differentiates specialised or energy kinesiology from similar but ineligible interventions (see Excluded therapies). The selected tools may include a wide range of interventions commonly considered to be ‘tools of the trade’ in kinesiology (e.g. acupressure, aromatherapy, myofascial interventions - see Background for other examples). However, it is expected that the selection of such interventions would be tailored to the individual based on the results of the manual muscle testing. As such, studies in which all participants are assigned to receive kinesiology plus acupressure, for example, will be excluded unless acupressure is a co-intervention in both arms (groups) of the study.

It is expected that the majority of studies will involve participants receiving kinesiology in individual sessions from a kinesiology practitioner. Except for the specific exclusions below, kinesiology interventions will be eligible irrespective of the mode of delivery (face-to-face or virtual), the training or qualifications of the teacher or practitioner (except if the training is Professional Applied Kinesiology™ – see Excluded therapies), the setting in which kinesiology is used, and the dose and duration of treatment. More details about each of these intervention features is considered in Section 3.3.2 Data extraction and Appendix 3.

Excluded therapies:

1. **Applied kinesiology (Professional Applied Kinesiology™)**² refers to the original form of kinesiology that is used only as a diagnostic tool, not a treatment modality [9, 34]. While PAK is used to inform decisions about treatment, resulting treatments are not considered to be part of PAK, instead falling within the scope of practice of the registered health professionals that use kinesiology for diagnosis (mainly chiropractors). Nor is there a premise in PAK that the body can heal itself if the person is aware of what is needed to facilitate healing (i.e. awareness of the root cause of a condition will facilitate a person's healing process). Practitioners of Professional Applied Kinesiology™ (PAK) must be certified by the International Board of Applied Kinesiology (IBAK). A prerequisite for PAK certification is completion of a tertiary qualification that fulfils requirements for registration in a relevant health profession (e.g. chiropractic, osteopathy, physiotherapy or medicine). [9, 33, 34]
2. **Kinesiology** as used to refer to human movement science in North America (especially in Canada), and related terms for practitioners of human movement science.
3. **Kinesiology taping** as in a specific method used mainly by physiotherapists.
4. **Kinesiotherapy** which is sometimes used to refer to “the main, active component of physiotherapy and a means of restoring the range of movement, improving muscle strength and endurance, improving movement coordination, increasing the aerobic capacity and inducing a global sensation of wellness” [35].
5. **Other systems that use kinesiology muscle testing** (e.g. Total Body Modification, PSYCH-K, Body talk).

Comparisons

1. kinesiology *versus* any inactive comparator (placebo/sham, no intervention, wait list control, usual care).
2. kinesiology *versus* evidence-based gold standard treatment(s) (see below for selection method)
3. kinesiology *versus* other active comparators (for inclusion in evidence inventory only, not the synthesis – See below)

Comparisons 1 and 2 will be addressed in separate syntheses (meta-analyses). Each synthesis will address a broad outcome domain with studies grouped within by population group (where appropriate; see Figure 2 for examples). Where a study includes multiple arms, with at least one eligible comparator (e.g. a placebo control arm), we will include the eligible comparison(s). Co-interventions are eligible providing the same co-intervention is in both arms of the comparison.

For comparison 2, evidence-based gold standard treatments will be identified through the prioritisation process (Figure 1, step 3). We will provide the NTWC and NTREAP with a list of active comparators identified from included studies (step 2). Studies with active comparators will not contribute to the synthesis except in the exceptional circumstance where the NTWC considers that the comparator intervention is an accepted, evidence-based ‘gold standard’ of care for the population in the studies, and there are studies suitable for conducting a synthesis (meta-analysis)

² PAK is trademarked so that it can only be used by certified members of International College of Applied Kinesiology (ICAK) <https://www.icaka.org.au/Applied-Kinesiology-Certification>

(i.e. comparable PICO criteria, low risk of bias). These judgements will be made blinded to the studies and study results, to the fullest extent possible. For studies involving other active comparators, we will provide an inventory of available evidence, tabulating a brief description of the characteristics of the Population, Intervention, Comparator, and Outcomes (PICO) for each study.

Exclusions. In line with the main review objective, which is to examine the effects of kinesiology rather than the comparative effects of different implementations of kinesiology, we will exclude head-to-head comparisons of kinesiology from the review (see exceptions below). For example, we will exclude studies where the only comparator is:

- a different dose (frequency, duration, schedule or combination thereof) of the same kinesiology system (e.g. different numbers of sessions)
- a different mode of delivery of the same kinesiology system (e.g. face-to-face versus online),
- where the person delivering kinesiology has a different qualification, or level of experience (e.g. accredited kinesiologist versus other health professional with some kinesiology training),
- or combinations of the above.

3.1.4 Types of outcomes

Any study that measures a health outcome will be eligible for the review, however the synthesis will be limited to outcomes of importance for decision-makers as identified through the process in Figure 1 and detailed below. Outcomes eligible for this review are those that align with the reasons why kinesiology is sought by patients and prescribed by practitioners. In principle, this may include any patient-important outcome that helps elucidate the effects of kinesiology on an underlying condition or its symptoms, recovery, rehabilitation, or prevention of disease among people with specific risk factors or pre-conditions.

Example outcome domains are shown in Figure 2. The example outcome domains are intended to illustrate the breadth of outcomes likely to be important for understanding the effects of kinesiology across conditions, as identified from the background literature on kinesiology.

The grouping of related outcomes within each domain (Figure 2) is based on ICD11 codes and the COMET outcome taxonomy[26, 27]. These systems provide a widely agreed and understood structure for categorising different outcomes

Prioritisation and selection of outcomes for summary and synthesis

Outcome prioritisation. To accommodate the breadth of relevant outcomes, the outcome domains and population-specific outcomes for inclusion in the synthesis will be determined through the prioritisation process (Figure 1).

To prioritise the most important outcomes for this review:

- We will compile a list of specific outcomes from included studies and example outcome measures (without results or identification of studies).
- Outcomes in the list will be categorised by the outcome domains and population groups in Figure 2. Outcomes that fall outside the proposed outcome domain will also be listed.

- The NTWC will be asked to indicate whether each of the listed outcome domains (or specific outcomes) is critical, important or of limited importance for understanding the effects of kinesiology on each population group. Only critical and important outcomes will be considered in the review.

Outcome selection. From each study, we will select only one outcome per outcome domain for data extraction (results), risk of bias assessment and inclusion in the summary and synthesis. The exception is for timeframe of measurement (see below).

For each outcome domain, we anticipate that there may be considerable multiplicity of results arising as follows.

- (1) Across studies, both the specific outcomes and methods used to measure each outcome will vary.
- (2) Within studies, results may be reported for multiple outcomes within a domain (e.g. pain intensity overall, pain on walking), multiple measures (e.g. visual analogue scale; overall and subscale scores from the Western Ontario and McMaster Universities Osteoarthritis Index), at multiple timepoint, or combinations of all three.

We will address this by using the following approach to select one outcome per domain from each study and by using the standardised mean difference (SMD) to quantify the effects of kinesiology. Using the SMD enables results to be combined for meta-analysis irrespective of the measure used, thus ensuring that any study that reports an outcome within a domain can be included in the analysis (see 3.3.9).

- An initial hierarchy of population-specific outcomes and measures will be presented to the NTWC for discussion and approval (e.g., a hierarchy of pain outcomes and measures for fatigue).
- Where possible, the initial hierarchy will be based on outcome hierarchies used in published Cochrane reviews, systematic reviews of measures that provide evidence of the relevance and validity of measures, and core outcome sets.
- We will also seek advice on the most relevant time point for outcome measurement. This is likely to be at the end of the intervention period
- We will follow the agreed hierarchy of population-specific outcomes measures and timepoints to select the most relevant and valid measure of each outcome domain available from each study for inclusion in the synthesis.

Exclusions:

- experience of care (e.g., satisfaction),
- safety,
- quality, and
- economic outcomes.

Studies will not be excluded from the review based on outcome, except where it is possible to confirm unequivocally (e.g., from a registry record or a priori protocol) that the study only measured experience of care, safety, quality, or economic outcomes.

3.2 Search methods for identification of studies

3.2.1 Electronic searches

We will conduct searches of the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, Embase, Emcare, AMED, CINAHL and Europe PMC (for preprints). Details of completed or ongoing studies will be sought from ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform. Searches will not be limited by date, language, or format of publication.

3.2.2 Selection of search terms

Although all the databases have a subject heading for kinesiology, the term is not included in the search strategy since the definition covers a form of kinesiology (applied kinesiology, muscle testing) that is different to the specialised or energy kinesiology that is in scope for this review.

For textwords we explored related concepts, such as "manual muscle testing" and kinesiotherapy, but both are more commonly used in settings and by health professionals distinct from kinesiology as a natural therapy. We will limit retrieval of irrelevant records (e.g. kinesiio taping studies, kinesiotherapy) by restricting the textword search accordingly.

The term kinesiology is combined with a very broad study design filter, based partly on a validated filter for non-randomised studies by Waffenschmidt et al [36] but with fewer restrictions and additional textwords, plus the floating subheading 'therapy' in the MEDLINE and Embase search. (See Appendix 1 for proposed search strategies.)

3.3.3 Searching other resources

The 2015 overview of kinesiology identified three systematic reviews that included studies of kinesiology. The three RCTs and 10 non-randomised studies from these reviews will be added to the records we screen. We will also look at studies included in any systematic reviews of kinesiology published since 2014 that our searches identify.

We will screen studies provided by the public and key stakeholders (via Department of Health), NTREAP and NTWC for eligibility. Where these groups recommend particular systematic reviews, we will examine the references to identify potentially eligible studies.

We will examine the reference lists of all included studies to identify additional trials or studies, and conduct forwards citation searching (i.e. looking for studies that have cited included studies) using Citation Chaser (<https://estech.shinyapps.io/citationchaser/>).

For studies included in the review, we will conduct supplementary searches to check whether any are listed as fraudulent or retracted, or have errata and comments associated with them. This will involve retrieving the PubMed record linked to each study and checking for the following Publication Types: Retracted Publication, Retraction of Publication, Expression of Concern, Published Erratum, and Comment. For reports of studies not indexed in PubMed, we will check the source database. We will also search the Retraction Watch Database (<http://retractiondatabase.org/>) using the term 'kinesiology'. As appropriate, we will consult Chapter 4 (section 4.4.6) of the Cochrane Handbook and section 3.9 of Technical Supplement for additional guidance.

3.3 Data collection and analysis

3.3.1 Selection of studies

Records from the database and register searches will be imported into EndNote and duplicates removed. All remaining records will be imported into Covidence for screening. Records submitted through the Department, NTREAP or the Committee will be screened to confirm that the type of study is eligible, then non-duplicate records will be imported into Covidence for screening alongside other studies.

We will pilot guidance for title and abstract screening on a sample of 50 records to ensure the eligibility criteria are being applied consistently by three reviewers (SB, MM, SM). If needed, we will amend the screening guidance (but not the eligibility criteria) to enhance consistency. All records at title and abstract stage will be independently screened by at least two reviewers. All records selected for full-text screening will be reviewed independently by two reviewers. Disagreements at either stage of screening will be resolved by consensus among members of the review team. Where disagreement cannot be resolved, advice will be sought from the NTWC (which will be provided with PICO characteristics for the de-identified study).

Studies confirmed as meeting the eligibility criteria, but for which results are not available in a published report, will be included in a list of 'ongoing studies'.

The following will be included in a list of 'studies awaiting classification'.

- Studies that are only published as abstracts or for which a full report is not available (i.e. we will not seek further information from study authors to confirm eligibility).
- Studies identified by, or submitted to, the review team after the date of the last search.
- Studies confirmed as likely to be eligible, but for which no English language translation of the full-text publication is available. Studies for which eligibility cannot be confirmed following translation of the title and abstract using Google translate, will be listed separately (Figure 3).

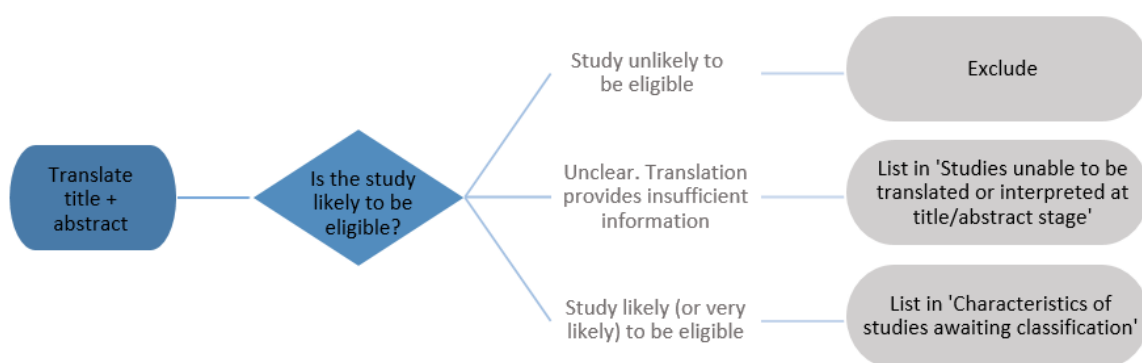


Figure 3. Flowchart showing handling of studies in languages other than English (reproduced from NHMRC framework for natural therapies systematic reviews [37]).

Studies that do not meet the eligibility criteria will be excluded and the reason for exclusion will be recorded at full-text screening. These studies will be included in a 'Characteristics of excluded studies' table in which the reason for exclusion is reported.

The search and study selection steps will be summarised in a PRISMA flow diagram.

For studies that originated from the call for evidence, NTREAP, or the Committee, we will record and report exclusion decisions irrespective of whether the study was excluded during title and abstract screening or full text review. We will document the flow of these studies through the review in the PRISMA flow chart and annotate tables with the source.

Dealing with duplicate and companion publications

Multiple publications to the same study (e.g. protocols, trial registry entries, trial reports) will be identified and linked prior to the data extraction stage in Excel. Each study will be given a unique identifier and all linked records cited in the final report. Records will be matched using trial registry numbers. Where these are not available, we will consider author names, trial name, trial location(s) and number of participants.

3.3.2 Data extraction and management

Three authors (MM, SB or SM) will pre-test the data extraction and coding form on 3-5 studies (as needed to achieve consistent coding) using REDCap electronic data capture tools hosted at Monash University [38, 39]. These studies will be purposefully selected from the included studies to cover the diversity of data types anticipated in the review. One author (SB) will review the extracted and coded pilot data for completeness, accuracy and consistency. Where needed, advice will be sought from a clinical advisor and biostatistician (JM) to ensure data are extracted as planned. Revisions to the data extraction form and guidance will be made as required to maximise the quality and consistency of data collection.

For each included study, one review author (MM, SB, or SM) will extract study characteristics and quantitative data using a pre-tested data extraction and coding form, with a 10% random sample extracted by a second author (with further sampling if needed until 80% agreement is achieved). For studies extracted by a single author, a second author (MM, SB, or SM) will independently verify the quantitative data. Discrepancies will be resolved through discussion, and advice sought from a third reviewer, the biostatistician (JM), if agreement cannot be reached or for more complex scenarios. NTWC will be asked to adjudicate on borderline decisions (e.g. whether an otherwise healthy population has ‘risk factors’).

We will extract information relating to the characteristics of included studies and results as follows.

1. Study identifiers and characteristics of the study design

- Study references (multiple publications arising from the same study will be matched to an index reference; code as index paper, protocol, registry entry, results paper 1, 2, ...)
- Study name, location, commencement date, and trial registration number
- Study design (categorised as randomised trial [‘individually randomised’, ‘cluster randomised’, ‘cross-over’, or ‘other’] or non-randomised study of intervention [‘non-randomised controlled trial’, ‘prospective cohort study’, ‘retrospective cohort study’, ‘controlled interrupted time series study’, or ‘controlled before-after study’]).
- Funding sources and funder involvement in study.
- Financial and non-financial interests declared by investigators.

2. Study design

- Study design label: coded according to the design identified after applying the checklist of design features (Appendix 2)
 - Detailed coding of design features of each study using the checklist of design features (Appendix 2)
3. Characteristics of each intervention group (including comparator groups)
- Characteristics of the intervention structured by domains of the Template for Intervention Description and Replication (TIDIER) checklist [40] (see Appendix 3 for TIDIER domains, codes and an example of coding for kinesiology).
 - Number of participants: randomised to each group, at follow up for selected outcome, and included in analysis and reasons for loss to follow-up
 - For NRSIs, details of the methods used to collect information on the intervention allocated or received by each participant (i.e. the information used to classify intervention status)
4. Characteristics of participants
- Participant eligibility criteria (verbatim or precis)
 - Age (mean, range)
 - Sex
 - Population group: coded using categories specified in the final analytic framework for the review (e.g. chronic musculoskeletal pain (low back pain/neck pain/arthritis), other chronic pain, movement disorders (Parkinsonism), injury and falls prevention, stress, anxiety and mood disorders, pregnancy, chronic respiratory conditions)
 - Condition: specific underlying condition as described in study (e.g. asthma, depression), ICD11 code, information about severity (if relevant)
 - Other characteristics of importance within the context of each study
5. Outcomes assessed and results
- Outcomes measured (a list of all outcomes, categorised according to the broad domains specified in the final analytic framework for the review, or as 'other' if none of the outcome domains apply)
 - For outcomes selected for inclusion in the summary and synthesis of results:
 - Outcome domain: categorised according to the broad domains specified in the final analytic framework for the review (e.g. pain, physical function (disability), fatigue, sleep quality, health-related quality of life, emotional functioning / wellbeing, condition-specific symptoms)
 - Outcome as described in the included study (verbatim or precis)
 - Measurement method (e.g. Asthma Control Questionnaire (ACQ) or Asthma Control Test (ACT)), information required to interpret measures (scale range and direction, minimally important difference) and time point (exact, and time-frame categorised as 'immediate' or 'longest follow-up')

- Results including: summary statistics by group (means and standard deviations, or number of events for cognitive outcomes that have been dichotomised, and sample size), estimates of intervention effect (e.g. mean differences (or adjusted mean differences for NRSIs), confidence intervals, t-values, p-values, or risk ratios/odds ratios for binary outcomes). For adjusted estimates, we will extract information on the analysis method, how confounding was adjusted, and which confounders were adjusted for (applies primarily to NRSIs).
- Data required to support risk of bias judgements (see Section 3.3.3) [41-44]

3.3.3 Assessment of risk of bias of included studies

Assessment of risk of bias in RCTs

We will assess the risk of bias in included studies using the revised Cochrane ‘Risk of bias’ tool (RoB 2) for randomised trials [41, 44] or the ROBINS-I tool for non-randomised studies of interventions [31, 42, 43]. The assessment will be made for each critical (or important) outcome included in the synthesis. Our assessment will be based on the effect of assignment to the intervention.

RoB 2 addresses five domains:

1. bias arising from the randomisation process;
2. bias due to deviations from intended interventions;
3. bias due to missing outcome data;
4. bias in measurement of the outcome;
5. bias in selection of the reported result.

For cluster trials and cross-over trials, we will use the variants of the RoB 2 tool specific for each design [30, 41]. Non-randomised clinical trials that do not include the study design features that we have specified for a ‘randomised’ trial will be assessed with ROBINS-I.

ROBINS-I addresses seven domains:

1. bias due to confounding
2. bias in selection of participants into the study
3. bias in classification of interventions
4. bias due to deviations from the intended intervention
5. bias due to missing data
6. bias in measurement of outcomes
7. bias in the selection of the reported result

In assessing risk of bias due to confounding (domain 1), we will answer the optional question about the predicted direction of bias (to determine whether residual confounding favours the control, thus lessening concerns that any observed beneficial effects are due to bias). In our overall judgement for domains 1 and 2, we will consider whether large effects or a dose response

mitigate concerns about bias due to confounding or selection of participants into the study [43, 45].

To promote concordance, the assessment will be piloted by three review authors (MM, SB, SMC) on 3-5 studies until consistent judgments are achieved across a range of scenarios. One review author (MM, SB or SMC) will then apply the tool to the selected results from each study following the RoB 2 or ROBINS-I guidance [41, 43], and a second author will verify the assessments (SB or SMC). Supporting information and justifications for judgements for each domain (i.e. for ROB 2: low, some concerns, high risk of bias; for ROBINS-I: low, moderate, serious, critical risk of bias) will be recorded. We will derive an overall summary of the risk of bias from each assessment, following the algorithm in the RoB 2 guidance [41] and in a forthcoming revision of guidance to ROBINS-I. Disagreement between review authors will be resolved through discussion, and a third review author (SB, SM or JM) will adjudicate where agreement cannot be reached.

For NRSIs, if we judge a result to be at “critical” risk of bias on the first domain (bias due to confounding), we will not assess other domains, since the overall risk of bias for the result would be “critical” by default. Results judged to be at “critical” risk of bias overall will be excluded from the summary and syntheses of results, and will not contribute to our conclusions.

Selection of effects for ROB assessment and synthesis. When multiple effects of the intervention using different approaches are presented in a study report, we will select one effect for inclusion in the meta-analysis and for RoB assessments. For randomised trials, the selected effect will be chosen according to the following hierarchy, which orders the approaches from (likely) least to most biased for estimating the *effect of assignment to the intervention*: 1. the effect that corresponds to a full intention-to-treat analysis, where missing data have been multiply imputed, or a model-based approach has been used (e.g. likelihood-based analysis, inverse-probability weighting); 2. the effect corresponding to an analysis that adheres to intention-to-treat principles except that the missing outcome data are excluded; 3. the effect that corresponds to a full intention-to-treat analysis, where missing data have been imputed using methods that treat the imputed data as if they were observed (e.g. last observation carried forward, mean imputation, regression imputation, stochastic imputation); or 4. the effect that corresponds to an 'as-treated' or 'per-protocol' analysis, or an analysis from which eligible trial participants are excluded [41, 44]. For NRSIs, the selected effect will be the estimate adjusted for critical confounding domains.

Pre-specification of confounding factors and co-exposures

Confounding domains are “prognostic variables (factors that predict the outcome of interest)” that also predict whether an individual receives one or the other interventions of interest [31]. ROBINS-I defines important confounding domains as those “for which, in the context of [a specific] study, adjustment is expected to lead to a clinically important change in the estimated effect of the [exposure]”. It is recommended that users applying ROBINS-I should consider in advance the confounding factors and co-interventions that have the potential to lead to bias in included studies. Due to the diversity of conditions and outcomes eligible for this review, confounding domains and co-exposures will be identified after the prioritisation process (Figure 1) but prior to assessment of studies.

3.3.4 Measures of treatment effect

We anticipate that many of the outcomes will be continuous (e.g. pain, physical function), and that varying measurement instruments will be used to measure the same underlying construct across the studies. For this reason, we will quantify the effects of kinesiology using the standardised mean difference (SMD) (implementing the Hedges’ adjusted *g* version). In trials where a continuous measure has been dichotomised (e.g. a continuous pain scale is dichotomised into

improvement or no improvement) and analysed as binary outcomes, we will re-express reported, or calculated, odds ratios as SMDs [46]. For dichotomous (e.g. experienced a fall), count (e.g. number of falls in a period of time) and time-to-event (e.g. time to first fall) outcomes, we will quantify the effects of kinesiology using risk ratios (RR), rate ratios, and hazard ratios, respectively, where possible.

Thresholds for interpreting the effects. Following the GRADE minimally contextualised approach for interpreting findings, all interpretations will be based on where the combined estimate of effect (point estimate) lies in relation to a pre-specified threshold for an important effect (i.e. whether an important effect is observed or not) and the direction of effect (beneficial or harmful) [47]. Given the wide range of conditions and outcomes in this review, it is not possible to specify specific thresholds for interpreting the size of the effect for each outcome. We plan to use Cohen's guiding rules for interpreting SMDs where 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 a large effect [48]. In practice, our interpretation will be based on whether there was an important effect or not [47, 49, 50], with an SMD of 0.2 standard units set as the threshold for an important difference. If the SMD falls within the prespecified range of -0.2 to 0.2 (i.e. within both thresholds), the effect of Kinesiology will be considered to be no different from control. An SMD above 0.2 or below -0.2 will be interpreted as an important effect. Where a valid and reliable minimal important difference (MID) is available for a familiar measure of relevance to the population groups in the meta-analysis, we will re-express the SMD in units of the measure and interpret the effect in relation to the MID if feasible to do so [48]. For dichotomous outcomes, we will use a relative risk increase of 25% (or above) and a relative risk reduction of 20% as the thresholds for an important effect. We will also express relative effects in absolute terms, and consider whether the absolute risk difference with intervention still indicates an important effect, seeking advice from the NTWC if decisions are borderline.

3.3.5 Unit of analysis issues

In this review, unit of analysis issues may arise from non-standard designs (cluster trials, cross-over trials) or from trials with more than two eligible intervention groups. In the following we outline the methods for making adjustments when necessary. Any adjustments will be documented (e.g. assumed intra-cluster correlation and average cluster size). We will also report when necessary adjustments were unable to be made due to missing information.

For cluster randomised trials that have not appropriately accounted for correlation in observations within clusters, we will attempt a re-analysis. We will do this by inflating the variance of the intervention estimates by a design effect (DEFF). The DEFF is calculated from two quantities – an intra-cluster correlation (ICC) and the average cluster size. Estimates of ICC will be imputed from other cluster trials included in the review, where possible, or by using external estimates from empirical research (e.g. Bell 2013 [51]). The average cluster size will be calculated from reported information in the trial.

For cross-over trials where an appropriate paired analysis is not available, we will attempt to approximate a paired analysis by imputing missing statistics (e.g. correlation). Estimates of the missing statistics will be imputed from other cross-over trials included in the review, where possible, or by using external estimates from empirical research (e.g. Balk 2012 [52]).

For trials where more than one comparison from the same trial is eligible for inclusion in the same meta-analysis (e.g. instructors with different training in kinesiology; different numbers of lessons), we will combine intervention groups, where it makes sense to do so; otherwise, we will appropriately reduce the sample size so that the same participants do not contribute more than once.

3.3.6 Dealing with missing data

We will not contact trial authors to obtain missing information (e.g. study characteristics, description of conduct of the study) or aggregate level statistics (e.g. missing standard deviations). However, we will attempt to calculate statistics necessary for meta-analysis using algebraic manipulation of reported statistics (e.g. computing the standard error for the treatment effect from a reported p-value). When standard deviations cannot be calculated from available statistics, but interquartile ranges or ranges are reported, we will use the formula in Wan et al [53] to estimate approximate standard deviations. When neither of the above methods are possible, we will impute the standard deviation using the average standard deviation across trials included in the same meta-analysis that have used the same measurement tool. When means are missing, but medians are reported, we will use the formula in Wan et al [53] to estimate approximate means.

Our approach for dealing with missing outcome data within the primary trials will be through sensitivity analyses, where trials judged to be at a high risk of bias or some concerns will be excluded (Section 3.3.9 Data synthesis). Risk of bias ‘due to missing outcome data’ is considered within the overall bias judgement for a trial. A similar approach will be taken with NRSI, excluding studies judged to be at serious risk of bias.

3.3.7 Assessment of heterogeneity

We will assess statistical heterogeneity of the intervention effects visually by inspecting the overlap of confidence intervals on the forest plots, formally test for heterogeneity using the χ^2 test (using a significance level of $\alpha=0.1$), and quantify heterogeneity using the I^2 statistic [54].

3.3.8 Assessment of biases due to missing results

We will use a framework for assessing risk of bias due to missing results in which an assessment is made for each meta-analysis regarding the risk and potential impact of missing results from studies (termed ‘known-unknowns’) and the risk of missing studies (termed ‘unknown-unknowns’) [55]. We will use this framework to guide our assessments of whether there is ‘undetected’ or ‘suspected’ reporting bias for each of the comparisons in our GRADE assessment (Section 3.3.10 Summary of findings tables and assessment of certainty of the body of evidence).

In assessing ‘known-unknowns’, we will determine what studies meeting the inclusion criteria for the particular meta-analysis have missing results through examination of the publication’s methods section, trial registry entry (if available), and trial or study protocol (if available). We will make an assessment as to whether the missing result was potentially due the result itself (e.g. ‘not statistically significant’), and whether inclusion of the result could lead to a notable change in the meta-analysis (e.g. if the missing result is from a large trial). We will also assess the impact of missing results from studies reported in languages other than English that were judged as being likely to meet the eligibility criteria for each synthesis (see Section 3.1.1 Types of studies; Section 3.3.1 Selection of studies).

In assessing ‘unknown-unknowns’, we will judge whether the studies not identified were likely to have results eligible for inclusion (e.g. for broad outcome domains such as ‘pain’, it is likely that for particular conditions, missing studies would have been eligible for inclusion). We will use funnel plots and contour enhanced funnel plots to examine whether there is evidence of small study effects [56]. If there is funnel plot asymmetry, we will undertake a sensitivity analysis to compare the combined effect estimated from the random-effects model (primary analysis) with that estimated from a fixed (common) effect model. If the random-effects estimate is importantly larger than the fixed-effect estimate, with no explanation for the difference (e.g. differences in clinical populations or intensity of the delivery of intervention between small and large trials,

differences in risk of bias between small and large trials), then we will downgrade for ‘suspected’ reporting bias.

3.3.9 Data synthesis

Meta-analysis

Separate comparisons will be set up based on outcome domains agreed in the final framework (see Figure 2). These comparisons will be stratified in two ways. First, by the population groups in the final framework (see Figure 2 and 3.1.2 Types of participants for indicative groupings). This approach to structuring the meta-analysis will yield an overall estimate of the effect of kinesiology for the outcome (review objectives 1, 2 and 4), as well as estimates within each population group (review objective 2). Subgroup analysis by population group will allow examination of whether these population groups explain any observed statistical heterogeneity in the intervention effects (Section 3.3.9 Subgroup analysis). Second, by study design, where we will separately combine estimates from randomised trials and non-randomised studies. We will not calculate an overall estimate of the effect of kinesiology across these study design strata.

We will combine the effects using a random effects meta-analysis model, since we expect there to be clinical and methodological diversity across the trials that may lead to statistical heterogeneity. These analyses will use the restricted maximum likelihood estimator (REML) of between trial heterogeneity variance and the Hartung-Knapp-Sidik-Jonkman confidence interval method.

Forest plots will be used to visually depict the intervention effect estimates and their confidence intervals. Forest plots will be stratified by condition and risk of bias (within population group).

Summary and synthesis when meta-analysis is not possible

Available effect estimates (95% confidence intervals, p-values), details of scales (direction and range), risk of bias assessments, and intervention characteristics will be tabulated. Tables will be ordered by outcome domain, population group and risk of bias assessment.

For a particular comparison, if we are unable to analyse most of the effect estimates (due to incomplete reporting of effects and their variances, variability in the effect measures across the studies), we will consider alternative synthesis methods, such as calculating summary statistics of the effect estimates, combining p-values, or vote counting based on the direction of effect [57]. Our choice of method will be determined by the available data (e.g. summary statistics if data permit; other methods if the data are more limited).

Subgroup analysis and investigation of heterogeneity

We will undertake a subgroup analysis to examine whether population group explains any observed statistical heterogeneity in the intervention effects (see Figure 2 and 3.1.2 Types of participants for indicative population groupings).

Sensitivity analyses

We plan to undertake and report sensitivity analyses examining if the meta-analysis estimates are robust to the:

- *meta-analysis model*. In addition to fitting a random-effects model, we will fit fixed effect models. This analysis will be undertaken to investigate the impact of any small-study effects.

- *inclusion of trials judged to be at an overall high risk of bias or some concerns about risk of bias. We will exclude trials judged to be at an overall high risk of bias or some concerns.*

Results of the sensitivity analyses will be tabulated, including the meta-analysis estimate (and its confidence interval), along with details of the original and sensitivity analysis assumptions.

3.3.10 Summary of findings tables and assessment of certainty of the body of evidence

We will prepare GRADE summary of findings tables for each of the main comparisons, reporting results for critical and important outcome domains (up to seven). For each result, one author (SB) will use the GRADE approach to assess our confidence in where the effect lies relative to our threshold for a small effect (the certainty of evidence) (see 3.3.4 Measures of treatment effect). In accordance with detailed GRADE guidance [45, 48, 58], an overall GRADE of high, moderate, low or very low certainty will be reported for each result based on whether there are serious, very serious or no concerns in relation to each of the following domains.

1. **Risk of bias.** We will assess the overall risk of bias across all studies contributing to each synthesised result, considering the weight studies rated at high risk of bias and some concerns contribute to the analysis. The evidence will be rated down for serious concerns if most of the weight comes from studies at some risk of bias. Very serious or extremely serious concerns are more likely if studies at high risk of bias contribute considerable weight in the analysis and sensitivity analyses indicate that removing these studies changes the size of the effect (see 3.3.9 Sensitivity analyses).
2. **Inconsistency.** We will assess whether there is important, unexplained inconsistency in results across studies considering the overlap of confidence intervals (non-overlap indicating potentially important differences in direction or size of effect), statistical measures that quantify and test for heterogeneity (I^2 statistic, χ^2 test), results of subgroup analyses (see 3.3.7 Assessment of heterogeneity), and where the point estimates lie in relation to the threshold for an important effect (if all to one side of a threshold, we were less concerned) [59]. To enhance our interpretation of whether inconsistency is important, we will also examine the prediction interval (if sufficient data to calculate), considering whether it included values that lead to a different conclusion than an assessment based on the confidence interval [60]. Where a result is based on a single study, inconsistency will not be rated.
3. **Imprecision.** We will assess whether the confidence interval for each pooled effect estimate crosses our threshold for an important effect (e.g. including a small effect and little or no difference, which would lead to different interpretations) and, for large effects, whether the sample size meets the optimal information size (based on number of events for binary outcomes; sample size for continuous outcomes). In judging imprecision, we will use our thresholds for a small but important effect as specified in 3.3.4 Measures of treatment effect. We will rate down for serious imprecision when the confidence interval appreciably crosses one threshold, very serious when two thresholds are crossed, and extremely serious when the confidence interval is so wide that the estimate is considered uninterpretable [47].
4. **Indirectness.** We will assess whether there are important differences between the characteristics of studies included in each synthesis and the question we are seeking to address, such that the effects observed may not apply to our question (i.e. the applicability of the evidence). For example, differences between the delivery of kinesiology in the study compared to Australia that are likely to influence the size of effect.

5. **Publication bias.** Our judgement of suspected publication bias will be based on assessment of bias due to missing results (see 3.3.8). In these assessments, we will also consider the potential impact on each synthesised result of excluding studies in languages other than English.
6. **Upgrading domains** (large effect size, dose response gradient, opposing plausible residual confounding). For NRSIs, we will consider large effects, any dose response, and opposing residual confounding as part of the ROBINS-I assessment (judging whether any of these factors mitigate concerns that observed effects are due to residual confounding) [45]. For this reason, upgrading of NRSIs will not be considered. There is no precedent for rating up the evidence from randomised trials, however in principle, these domains apply to any body of evidence so are included here for completeness.

Using GRADE decision rules, we will derive an overall GRADE for the certainty of evidence for each result included in the summary of findings table [21, 58]. A result from a body of evidence comprised of randomised trials begins as ‘high’ certainty evidence (score=4), and can be rated down (-1 or -2) for serious or very serious concerns on any GRADE domain that reduces confidence that kinesiology has at least a small effect (as determined by the pre-specified thresholds) [21, 48, 58]. A result from a body of evidence comprised of NRSIs will also begin as ‘high’ certainty, with the same rules for rating down used as for randomised trials. The option to rate down by -3 may be considered if there are extremely serious concerns about risk of bias (especially for NRSIs) or imprecision (where results are compatible with both important benefit and important harm) [47, 61].

Summary of findings tables will be prepared using the GRADEpro GDT software (www.gradepro.org). The tables will include:

- estimates of the effects of kinesiology reported as standardised mean differences, and for binary outcomes relative and absolute effects
- the overall GRADE (rating of certainty) and an explanation of the reason(s) for rating down (or up) [62]
- the study design(s), number of studies and number of participants contributing data
- a plain language statement interpreting the evidence for each comparison and outcome, following GRADE guidance for writing informative statements [63].

Where results for a comparison and outcome from both randomised trials and NRSIs are available, these will be presented in the same row of the summary of findings table providing that the certainty of evidence available from randomised trials is not lowered by inclusion of NRSIs at higher risk of bias than the trials [64]. This approach will enable an overall conclusion to be drawn from the body of evidence where appropriate to do so. Doing so may also ameliorate concerns about indirectness (if direct evidence is available from NRSIs) and imprecision (if high or moderate certainty evidence is available from NRSIs). If the certainty of evidence is lowered by inclusion of NRSIs at higher risk of bias than RCTs, or if results from NRSIs and RCTs are not coherent (i.e. both leading to an interpretation of a small but important improvement), then we will report the results in separate rows with separate GRADEs and interpretations.

We will present the four levels of certainty of evidence in summary of findings tables with the following symbols and interpretations.

- High ($\oplus\oplus\oplus\oplus$): further research is very unlikely to change the confidence in the estimate of effect
- Moderate ($\oplus\oplus\oplus\ominus$): further research is likely to have an important impact in the confidence in the estimate of effect
- Low ($\oplus\oplus\ominus\ominus$): 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 ($\oplus\ominus\ominus\ominus$): any estimate of effect is very uncertain

4. References

1. National Health and Medical Research Council. Statement of requirement: Evidence evaluations for review of natural therapies (tranche two). 2020
2. About Touch For Health. [Internet]: International Kinesiology College. <https://www.ikc.global/about-touch-for-health/> (2020). Accessed 2023 Dec 11.
3. What is Touch For Health? [Internet] Nampa, ID: Touch for Health Kinesiology Association. <https://touchforhealth.us/about-touch-for-health/> (2023). Accessed 2023 Dec 11.
4. Hall S, Lewith G, Brien S, Little P. A review of the literature in applied and specialised kinesiology. *Complementary Medicine Research* 2008;15(1):40-6.
5. Grünenberg K, Hillersdal L, Walker HK, Boelsbjerg HB. Doing wholeness, producing subjects: Kinesiological sensemaking and energetic kinship. *Body & Society* 2013;19(4):92-119.
6. Steel A, McIntyre E, Harnett J, Foley H, Adams J, Sibbritt D, et al. Complementary medicine use in the Australian population: Results of a nationally-representative cross-sectional survey. *Sci Rep* 2018;8(1):17325. doi:10.1038/s41598-018-35508-y.
7. What is Kinesiology? [Internet] Kerrimuir, VIC: Australian Kinesiology Association Inc. <https://aka.asn.au/what-is-kinesiology/> (2023). Accessed 2023 Mar 14.
8. Jensen A. Muscle testing (kinesiology): panacea or placebo? *The Conversation*. 2012. [cited 2012 Dec 19]. <https://theconversation.com/muscle-testing-kinesiology-panacea-or-placebo-11075>.
9. Rosner AL, Cuthbert SC. Applied kinesiology: distinctions in its definition and interpretation. *J Bodyw Mov Ther* 2012;16(4):464-87. doi:10.1016/j.jbmt.2012.04.008.
10. Maine I, McIntyre M, Huynh P, Bullivant A. SUBMISSION: Natural Therapies Review Expert Advisory Panel 2019-20; Australian Government Department of Health; Citations for Published Scientific Research Studies; Tranche Two, Kinesiology. Australian Institute of Kinesiologists Ltd & Australian Kinesiology Association. 2020. <https://aka.asn.au/evidence-based-research/>. Accessed 2023 Mar 14
11. Inspire Freedom. Online Kinesiology Consultations. [Internet] Sydney, NSW. <https://www.inspirefreedom.com.au/online-kinesiologist/> (2024). Accessed 2024 Mar 18.
12. Kinesiology Session Structure. [Internet] Thirroul, NSW: Australian Institute of Kinesiologists Ltd. https://www.aik.org.au/experience_kinesiology.html (2022). Accessed 2023 Mar 14.
13. Kinesiology Training Pathways. [Internet] Kerrimuir, VIC: Australian Kinesiology Association Inc. <https://aka.asn.au/kinesiology-training/> (2023). Accessed 2023 Mar 14.
14. Study kinesiology. [Internet] Thirroul, NSW: Australian Insitute of Kinesiologists Ltd. https://www.aik.org.au/study_kinesiology.html (2022). Accessed 2023 Mar 14.
15. Australian Health Practitioner Regulatory Association. Public consultation on clearer regulation of medical practitioners who provide complementary and unconventional medicine and emerging treatments. Available from: <https://www.ahpra.gov.au/> Access date: 4 February 2021 2021;
16. Eardley S, Brien S, Little P, Prescott P, Lewith G. Professional kinesiology practice for chronic low back pain: single-blind, randomised controlled pilot study. *Complementary Medicine Research* 2013;20(3):180-8.
17. Steel A. PRACI Study: unpublished data. 2021.
18. Experience Kinesiology. [Internet] Thirroul, NSW: Australian Insitute of Kinesiologists Ltd. https://www.aik.org.au/experience_kinesiology.html (2022). Accessed 2023 Mar 14.
19. Australian Government Department of Health. Review of the Australian Government Rebate on Natural Therapies for Private Health Insurance. 2015.
20. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.1 (updated September 2020). Cochrane. Available from www.training.cochrane.org/handbook; 2020.
21. Schünemann HJ, Brozek J, Guyatt G, Oxman AD, editors. *Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach*. Hamilton, Canada: McMaster University; 2013.
22. Moher D, Shamseer L, Clarke M, Gherzi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews* 2015;4(1):1. doi:10.1186/2046-4053-4-1.

23. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *Bmj* 2015;349:g7647. doi:10.1136/bmj.g7647.
24. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj* 2021;372:n71. doi:10.1136/bmj.n71.
25. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *Bmj* 2021;372:n160. doi:10.1136/bmj.n160.
26. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *Journal of Clinical Epidemiology* 2018;96:84-92. doi:<https://doi.org/10.1016/j.jclinepi.2017.12.020>.
27. World Health Organisation. International statistical classification of diseases and related health problems (11th ed.) Available from: <https://www.who.int/classifications/classification-of-diseases> Access date: 15 January 2021 2019.
28. Reeves BC, Deeks JJ, Higgins JPT, Shea B, Tugwell P, Wells G. Chapter 24: Including non-randomized studies on intervention effects. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions version 63 (updated February 2022)* Cochrane, 2022 Available from www.trainingcochrane.org/handbook
29. Reeves BC, Wells GA, Waddington H. Quasi-experimental study designs series-paper 5: a checklist for classifying studies evaluating the effects on health interventions—a taxonomy without labels. *J Clin Epidemiol* 2017;89:30-42. doi:10.1016/j.jclinepi.2017.02.016.
30. Higgins JPT, Eldridge SM, Li T. Chapter 23: Including variants on randomized trials. . In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions version 61 (updated September 2020)* Cochrane Available from www.trainingcochrane.org/handbook; 2020.
31. Sterne JAC, Hernán MA, McAleenan A, Reeves BC, JPT. H. Chapter 25: Assessing risk of bias in a non-randomized study. In: Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor. *Cochrane Handbook for Systematic Reviews of Interventions version 63 (updated February 2022)* Cochrane, 2022 Available from www.trainingcochrane.org/handbook
32. McKenzie JE, Brennan SE, Ryan RE, Thomson HJ, Johnson RV, Thomas J. Chapter 3: Defining the criteria for including studies and how they will be grouped for the synthesis. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Welch V, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd ed. Chichester (UK): John Wiley & Sons; 2019.
33. National Health and Medical Research Council. Kinesiology description developed in conversation with the National Health and Medical Research Council's Natural Therapies Working Committee Chair and the Department of Health's Natural Therapies Review Expert Advisory Panel (February 2020). 2020
34. International College of Applied Kinesiology (Australian Chapter). Professional Applied Kinesiology Certification Series. www.icaka.org.au Accessed 8 Nov 2023. <https://www.icaka.org.au/Applied-Kinesiology-Certification> (2023). Accessed 2023.
35. *Dictionary of Rheumatology*. Switzerland: Springer International Publishing; 2016.
36. Waffenschmidt S, Navarro-Ruan T, Hobson N, Hausner E, Sauerland S, Haynes RB. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. *Research Synthesis Methods* 2020;11(5):617-26.
37. National Health and Medical Research Council. Draft framework for protocols for systematic reviews of randomised controlled trials and non-randomised studies of interventions. 2020.
38. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics* 2019;95:103208. doi:<https://doi.org/10.1016/j.jbi.2019.103208>.
39. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics* 2009;42(2):377-81. doi:<https://doi.org/10.1016/j.jbi.2008.08.010>.
40. Hoffmann T, Glasziou P, Barbour V, Macdonald H. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj* 2014;1687:1 - 13.

41. Higgins JPT, Savovic J, Page MJ, Sterne JAC. Revised Cochrane risk-of-bias tool for randomized trials (RoB 2). Available from <https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2>. Access date: 8 February 2021. 2019;
42. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *Bmj* 2016;355:i4919. doi:10.1136/bmj.i4919.
43. Sterne JAC, Higgins JPT, Elbers RG, Reeves BC, and the development group for ROBINS-I. Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I): detailed guidance, updated 12 October 2016. Available from <http://www.riskofbias.info> [accessed 7 September 2022]. 2016.
44. 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. doi:10.1136/bmj.l4898.
45. 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. doi:10.1016/j.jclinepi.2018.01.012.
46. Higgins JPT, Li T, (editors) DJ. 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 63 (updated February 2022)* Cochrane, 2022 Available from www.trainingcochrane.org/handbook
47. Zeng L, Brignardello-Petersen R, Hultcrantz M, Siemieniuk RAC, Santesso N, Traversy G, et al. GRADE guidelines 32: GRADE offers guidance on choosing targets of GRADE certainty of evidence ratings. *J Clin Epidemiol* 2021;137:163-75. doi:10.1016/j.jclinepi.2021.03.026.
48. Schünemann HJ, Vist GE, Higgins J, Santesso N, Deeks JJ, Glasziou P, et al. Chapter 15: Interpreting results and drawing conclusions. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Welch V, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd ed. Chichester (UK): John Wiley & Sons; 2019.
49. Zeng L, Brignardello-Petersen R, Hultcrantz M, Mustafa RA, Murad MH, Iorio A, et al. GRADE Guidance 34: update on rating imprecision using a minimally contextualized approach. *J Clin Epidemiol* 2022;150:216-24. doi:10.1016/j.jclinepi.2022.07.014.
50. Brennan SE, Johnston RV. Research Note: Interpreting findings of a systematic review using GRADE methods. *Journal of Physiotherapy* 2023. doi:<https://doi.org/10.1016/j.jphys.2023.05.012>.
51. Bell ML, McKenzie JE. Designing psycho-oncology randomised trials and cluster randomised trials: variance components and intra-cluster correlation of commonly used psychosocial measures. *Psychooncology* 2013;22(8):1738-47. doi:10.1002/pon.3205.
52. Balk EM, Earley A, Patel K, Trikalinos TA, Dahabreh IJ. Empirical Assessment of Within-Arm Correlation Imputation in Trials of Continuous Outcomes. *Methods Research Report*. (Prepared by the Tufts Evidence-based Practice Center under Contract No. 290-2007-10055-I.) AHRQ Publication No. 12(13)-EHC141-EF. Rockville, MD: Agency for Healthcare Research and Quality. November 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm. In; 2012.
53. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014;14:135. doi:10.1186/1471-2288-14-135.
54. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21(11):1539-58. doi:10.1002/sim.1186.
55. Page MJ, Higgins JPT, Sterne JA. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions version 61 (updated September 2020)* Cochrane Available from www.trainingcochrane.org/handbook
56. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 2008;61(10):991-6. doi:10.1016/j.jclinepi.2007.11.010.
57. McKenzie JE, Brennan SE. Chapter 12: Synthesizing and presenting findings using other methods. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Welch V, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd ed. Chichester (UK): John Wiley & Sons; 2019.

58. Schünemann HJ, Higgins J, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Welch V, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd ed. Chichester (UK): John Wiley & Sons; 2019.
59. Guyatt G, Zhao Y, Mayer M, Briel M, Mustafa R, Izcovich A, et al. GRADE Guidance 36: Updates to GRADE's approach to addressing inconsistency. *J Clin Epidemiol* 2023. doi:10.1016/j.jclinepi.2023.03.003.
60. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, et al. CINeMA: An approach for assessing confidence in the results of a network meta-analysis. *PLOS Medicine* 2020;17(4):e1003082. doi:10.1371/journal.pmed.1003082.
61. Piggott T, Morgan RL, Cuello-Garcia CA, Santesso N, Mustafa RA, Meerpohl JJ, et al. Grading of Recommendations Assessment, Development, and Evaluations (GRADE) notes: extremely serious, GRADE's terminology for rating down by three levels. *Journal of Clinical Epidemiology* 2020;120:116-20. doi:10.1016/j.jclinepi.2019.11.019.
62. Santesso N, Carrasco-Labra A, Langendam M, Brignardello-Petersen R, Mustafa RA, Heus P, et al. Improving GRADE Evidence Tables part 3: Guidance for useful GRADE certainty in the evidence judgments through explanatory footnotes. *J Clin Epidemiol* 2016. doi:10.1016/j.jclinepi.2015.12.006.
63. Santesso N, Glenton C, Dahm P, Garner P, Akl EA, Alper B, et al. GRADE guidelines: 26. informative statements to communicate the findings of systematic reviews of interventions. *Journal of Clinical Epidemiology* 2019. doi:<https://doi.org/10.1016/j.jclinepi.2019.10.014>.
64. Cuello-Garcia CA, Santesso N, Morgan RL, Verbeek J, Thayer K, Ansari MT, et al. GRADE guidance 24 optimizing the integration of randomized and non-randomized studies of interventions in evidence syntheses and health guidelines. *J Clin Epidemiol* 2022;142:200-8. doi:10.1016/j.jclinepi.2021.11.026.

Appendix 1: Database and register search strategies

Cochrane Central Register of Controlled Trials

#	Search strategy
1	(kinesiolog* not (kinesio* tap* or kinesiotap* or kinesio therapy or kinesiotherapy)):ti,ab,kw

MEDLINE via Ovid

#	Search strategy
1	(kinesiolog* not (kinesio* tap* or kinesiotap* or kinesio therapy or kinesiotherapy)).ti,ab,kf.
2	(Clinical Study or Comparative Study or Evaluation Study).pt. or exp Epidemiologic Studies/ or (trial or study or cohort or control or experimental or group* or systematic review).ti,ab,kf. or th.fs.
3	1 and 2

Embase via Ovid

#	Search strategy
1	(kinesiolog* not (kinesio* tap* or kinesiotap* or kinesio therapy or kinesiotherapy)).ti,ab,kf.
2	exp Controlled Study/ or (trial or study or cohort or control or experimental or group* or systematic review).ti,ab,kf. or th.fs.
3	1 and 2

AMED via Ovid

#	Search strategy
1	(kinesiolog* not (kinesio* tap* or kinesiotap* or kinesio therapy or kinesiotherapy)).tw.
2	exp Research Design/ or exp Epidemiologic Methods/ or (trial or study or cohort or control or experimental or group* or systematic review).tw.
3	1 and 2

Emcare via Ovid

#	Search strategy
1	(kinesiolog* not (kinesio* tap* or kinesiotap* or kinesio therapy or kinesiotherapy)).ti,ab,kf.
2	exp Controlled Study/ or (trial or study or cohort or control or experimental or group* or systematic review).ti,ab,kf.
3	1 and 2

CINAHL

#	Search strategy
1	TI ((kinesiolog*) not (kinesio* tap* or kinesiolog* or kinesiolog* therapy or kinesiolog* therapy)) OR AB ((kinesiolog*) not (kinesio* tap* or kinesiolog* or kinesiolog* therapy or kinesiolog* therapy))
2	TX trial or study or cohort or control or experimental or group* or systematic review
3	S1 OR S2

Europe PMC

#	Search strategy
1	((TITLE:"kinesiology") OR (ABSTRACT:"kinesiology")) AND (SRC:"PPR")

ClinicalTrials.gov and WHO ICTRP

Kinesiology (Intervention/treatment field in ClinicalTrials.gov and Intervention field in WHO ICTRP)

Appendix 2: Study design features checklist

The following checklist is from Reeves et al [29]. The checklist is intended to be used as a tool for confirming eligibility of studies for the review and describing the design features of each included study. In this case, it is used to indicate how studies will be labelled in the review.

- Responses for each design were completed for this systematic review (not taken from the paper) to indicate the minimum features required for inclusion of each type of study (labelled 'Yes').
- Alternatives to these minimum features are labelled 'Possible'.
- In most cases, the combination of features labelled 'Yes' distinguishes the type of study (i.e. NRCT versus cohort).

Study design feature	Feature of				
	NRCT	Prospective cohort	Retrospective cohort	Controlled ITS	CBA
1. Was the intervention/comparator (answer 'yes' to more than one item, if applicable):					
• allocated to (provided for/administered to/chosen by) individuals?	Yes (or clusters)	Yes	Yes	Possibly	No
• allocated to (provided for/administered to/chosen by) clusters of individuals?	Possibly	No	No	Yes	Yes
• clustered in the way it was provided (by practitioner or organizational unit)?	Possibly	Possibly	Possibly	Possibly	Possibly
2. Were outcome data available (answer 'yes' to only one item):					
• after intervention / comparator only (same individuals)?	Yes	Yes	Yes	No	No
• after intervention/comparator only (not all same individuals)?	No	No	No	No	No
• before (once) AND after intervention/comparator (same individuals)?	Possibly	Possibly	Possibly	No	Possibly
• before (once) AND after intervention/comparator (not all same individuals)?	No	No	No	No	Yes
• multiple times before AND multiple times after intervention/comparator (same individuals)?	Possibly	Possibly	Possibly	Possibly	No (ITS)
• multiple times before AND multiple times after intervention/comparator (not all same individuals)?	No	No	No	Yes	No (ITS)
3. Was the intervention effect estimated by (answer 'yes' to only one item):					
• change over time (same individuals at different time-points)?	No	No	No	No	No
• change over time (not all same individuals at different time-points)?	No	No	No	No	No
• difference between groups (of individuals or clusters receiving either intervention or comparator)?	Yes	Yes	Yes	Yes	Yes
4. Did the researchers aim to control for confounding (design or analysis) (answer 'yes' to only one item):					
• using methods that control in principle for any confounding?	No	No	No	Yes	No
• using methods that control in principle for time invariant unobserved confounding?	No	No	No	No	Yes
• using methods that control only for confounding by observed covariates?	Yes [exclude if not]	Yes	Yes	Possibly	Possibly

Study design feature	Feature of				
	NRCT	Prospective cohort	Retrospective cohort	Controlled ITS	CBA
5. Were groups of individuals or clusters formed by (answer 'yes' to more than one item, if applicable):					
• <i>randomization?</i>	No	No	No	No	No
• <i>quasi-randomization?</i>	No	No	No	No	No
• some other action of researchers?	Yes	Possibly	Possibly	Possibly	No
• <i>time differences?</i>	No	No	No	Yes	No
• location differences?	Possibly	Possibly	Possibly	Possibly	Possibly
• healthcare decision makers/practitioners?	Possibly	Possibly	Possibly	Possibly	Yes
• participants' preferences?	Possibly	Yes	Yes	No	Possibly
• on the basis of outcome? e	No	No	No	No	No
• some other process? (specify)	Possibly	Possibly	Possibly	Possibly	Possibly
6. Were the following features of the study carried out after the study was designed (answer 'yes' to more than one item, if applicable):					
• characterization of individuals/clusters before intervention?	Yes	Yes	Possibly	Possibly	Possibly
• actions/choices leading to an individual/cluster becoming a member of a group?	Yes	Yes	Possibly	Possibly	Possibly
• assessment of outcomes?	Yes	Yes	Possibly	Possibly	Possibly
7. Were the following variables measured before intervention (answer 'yes' to more than one item, if applicable):					
• potential confounders?	Yes	Yes	Yes	Possibly	Possibly
• outcome variable(s)?	Possibly	Possibly	Possibly	Yes	Yes

NRCT: non-randomised controlled trial; ITS: interrupted time series; CBA controlled before-after studies

Appendix 3: TIDIER domains and example of application in kinesiology SR

Characteristic	Description	Standard headings (bolded) and codes illustrated with an example*
Brief name	Name or phrase	Professional Kinesiology Practice (PKP)
Why	Rationale, theory or goal of essential elements	Relief of chronic low back pain
What	Procedures, activities or processes including enabling/supporting activities Materials (physical or informational)	Back examination followed by full PKP treatment [not described] which comprised individualised techniques from the PKP repertoire at each session. Therapeutic conversation. Post-check measures of restriction and movement. Discuss changes with patient. Determine self-administered techniques for home maintenance. CO-INTERVENTIONS (code all that apply): written information leaflet, educational material, other [specify], none reported**
When and how much	Number of times delivered, over what period of time, number of sessions, their schedule, and their duration, intensity or dose	1 x treatment per week for 5 weeks (5 treatments total) [treatment duration not reported]
How	Modes of delivery ()	CODE (choose all that apply): individual sessions, group session; in person, online, other [specify]
Who provided	Intervention provider by category and their expertise, background and training	Administered by a provider: [Y/N] Self-administered: [Y/N] Provider (as identified by authors, code all that apply) <ul style="list-style-type: none"> <input type="radio"/> Kinesiology practitioner <input type="radio"/> Other natural therapist [specify] <input type="radio"/> Nurse (clinically qualified) <input type="radio"/> Allied Health [specify e.g. physiotherapist] <input type="radio"/> General practitioner/ physician [specify] <input type="radio"/> Research staff <input type="radio"/> Other [specify] <input type="radio"/> Not reported Highest level of training (for health-care providers only): postgraduate, bachelor, diploma, certificate, other [specify], not reported Trained in kinesiology? [Y/N/not reported] Details: Accredited with AKA
Where	Type of location where intervention occurred	Private kinesiology practice
Tailoring	If personalised, titrated or adapted describe what, why, when and how	n/a
Modifications	Not collected. Any modifications are likely to be part of protocol (i.e. tailoring)	
How well	Not collected. Used for questions about <i>adherence</i> to intervention (not assignment to intervention)	

** Codes highlighted in grey are the codes that apply to this example