HealthConsult

Better thinking. Better advice.

National Health and Medical Research Council

Whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice, for preventing and treating health conditions

Research Protocol

24 June 2021

HealthConsult Pty Ltd

ACN 118 337 821

Sydney 3/86 Liverpool Street, Sydney, NSW 2000

Phone (02) 9261 3707

Melbourne 429/838 Collins Street, Docklands, VIC 3008

Phone (03) 9081 1640

ACKNOWLEDGEMENTS

None.

AUTHORS

Lisa Fodero[1], Nerissa Soh[1], Katrin Schulz[1], Shari Stathis [1]

[1]HealthConsult Pty Ltd

Contact address: 3/86 Liverpool St, Sydney NSW 2000, Australia.
lisa.fodero@healthconsult.com.au

CONTRIBUTIONS OF AUTHORS

Dr Lisa Fodero: protocol design, review and revisions to draft protocol

Dr Nerissa Soh: protocol design, drafting of protocol, development of search strategies

Katrin Schulz: drafting of background

Shari Stathis: review and revision of draft protocol

DECLARATION OF INTEREST

The authors have no competing interests to declare.

SOURCES OF SUPPORT

Internal sources and funding: Office of the National Health and Medical Research Council (ONHMRC)

External sources: Natural Therapies Working Committee (NTWC)

Sponsor: ONHMRC and NTWC

Role of Sponsor: Process guidance as well as advice and input into project deliverables, including topic-specific advice on naturopathy

DISCLAIMER

This Research Protocol was prepared by HealthConsult on behalf of the National Health and Medical Research Council (NHMRC).

Table of Contents

Section Page

List of Tables

List of Figures

No table of figures entries found.

Abbreviations

AMED	Allied and Complementary Medicine Database
CAM	Complementary and alternative medicine
CINAHL	Cumulative Index of Nursing and Allied Health Literature
COMET	Core Outcome Measures in Effectiveness Trials
GRADE	Grading of Recommendations Assessment, Development and Evaluation
ISSG	Information Specialists' Sub-Group
ITT	Intention to treat
MECIR	Methodological Expectations of Cochrane Intervention Reviews
MeSH	Medical subject heading
NHMRC	National Health and Medical Research Council
NRSI	Non-randomised study of intervention
NTREAP	Natural Therapies Review Expert Advisory Panel
NTWC	Natural Therapies Working Committee ('the Committee')
ONHMRC	Office of the National Health and Medical Research Council
PBRN	Practice-Based Research Network
PICO	Population, Intervention, Comparator, Outcome
PRACI	Practitioner Research and Collaboration Initiative
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised controlled trial
ROBINS	Risk Of Bias In Non-randomised Studies of Intervention
SIGN	Scottish Intercollegiate Guidelines Network
TIDieR	Template for Intervention Description and Replication
WNF	World Naturopathic Federation

1 Background and Objectives of the Review

This Chapter sets the context for the Review as well as describes the intervention, the most common conditions of those who are undertaking naturopathic treatment, the importance and objectives of this Review.

1.1 CONTEXT

The National Health and Medical Research Council (NHMRC) has been engaged by the Department of Health to update the evidence underpinning the 2015 Review of the Australian Government Rebate on Natural Therapies for Private Health Insurance.

Naturopathy is one of seven natural therapies being reviewed in the first tranche of the work and one of the 16 therapies excluded from the private health insurance rebate as of 1st April 2019.

HealthConsult were engaged by the NHMRC on 4th March 2020 to review the clinical effectiveness of whole system, multi-modal, or single modal interventions delivered in the context of naturopathic practice. The purpose of the 2015 Review was to ensure that natural therapies are underpinned by a credible evidence base that demonstrates their clinical efficacy, cost-effectiveness and safety and quality. Further details regarding similarities and differences between this Review and the 2015 Review are described in Section 1.5.

This document presents the draft Research Protocol for a systematic review (the Review) into the effectiveness of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice. It presents the objectives, methods to locate, select and critically appraise studies and the method to collect and analyse data from the included studies.

1.2 DESCRIPTION OF THE CONDITIONS

The Review will include populations who are undertaking naturopathic treatment to either: prevent health condition/s (in at risk populations); or to treat, manage or delay progression of existing health conditions.

The Review is expected to include, but will not be limited to, the conditions identified by the Practitioner Research and Collaboration Initiative (PRACI). PRACI, which is the largest national practice-based research network (PBRN) for complementary healthcare (including naturopathy) in Australia[1] and has collected data on the most commonly self-reported conditions treated by naturopathic practitioners in Australia. The conditions which form >50% of those often seen by naturopaths include:[1]

- Fatigue (95% of respondents)
- Digestive Disorders (84%)
- Mental illness (77%)
- Irritable bowel syndrome (67%)
- Menstrual disorders (61%), and
- Insomnia/sleeping disorders (61%)

Conditions that were reported as sometimes seen, by >50% of naturopaths, include:[1]

- Hay fever (64% or respondents)
- Eczema/Psoriasis (57%)

- Headache/migraine (57%)
- Recurrent infections (54%), and
- Arthritis (51%).

1.3 DESCRIPTION OF THE INTERVENTION

For the purposes of this Review, the interventions of interest are whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice:

- **'Whole system'** in the context of naturopathy 'refers to the practice of naturopathic practice as a complex health care intervention that addresses simultaneously the multiple dimensions (physical, mental, spiritual, family, community, and environment) of an individual patient as pragmatically practised by naturopathic clinicians'.[2]
- **'Multi-modality'** refers to 'a minimum of two modalities as part of a single clinical approach to treatment of an individual'.[2]
- **'Single modality'** refers to the individual modalities used by a naturopath.

Naturopathy can be defined as a system of healthcare with a deep history of traditional philosophies and principles, utilising a number of natural therapy modalities to treat patients.[3] A naturopath typically sees patients via consultation in private clinical practice. An initial consultation is usually between 60 – 120 minutes duration[4] with follow-up consultations about 30 - 60 minutes.[5] In a typical consultation a naturopath takes a detailed case history and performs physical examinations such as pulse and tongue diagnosis, iridology and blood pressure.[4] A naturopath may also send a patient for laboratory testing (e.g. stool testing or pathology) to assist in determining a naturopathic diagnosis. Once a naturopathic diagnosis is confirmed, naturopaths usually develop a treatment plan using one or more modalities such as diet and lifestyle advice[4] or recommend other treatments like yoga and exercise.[5, 6] Naturopaths also provide maintenance for long term health[7], with some clients requiring follow-up appointments to refine treatment plans or maintenance appointments for a few months for chronic or ongoing conditions.[8]

The core naturopathic philosophies, principles and treatment modalities which form naturopathic practice and diagnosis are explained in more detail below.[9]

In Australia, 'Naturopathy' is not a regulated or registered profession. However, while not regulated, in order for naturopaths to obtain professional indemnity insurance, they need to be affiliated with a professional association. In recent years, naturopaths have developed an independent register for qualified naturopaths through the *Australian Register of Naturopaths and Herbalists* (ARONAH), which requires practitioners to meet 'competency standards' and have a minimum qualification.[10] Typical naturopathic training involves a diploma or degree level qualification[4] and some naturopathic organisations have minimum requirements in naturopathy such as an advanced diploma, a bachelor's degree, or another qualification in naturopathy or Western herbal medicine providing the practitioner can show evidence that they have been in regular practice in the last two to ten years.[11]

1.3.1 Core philosophies

Two core naturopathic philosophies are holism and vitalism. Holism (or holistic) refers to the 'whole' being greater than the sum of its 'parts'. In naturopathic practice this translates to treating both a health condition/ disease and an individual as a 'whole' not in isolation, and considers both internal (disease process) and external (environmental, social, cultural) factors that may contribute to the health of an individual.

Vitalism refers to the theory that every living organism has an innate 'vital force' or natural wisdom. To treat a condition using a 'vitalistic' approach is to encourage the body's natural ability to heal itself, rather than supressing or masking symptoms *(e.g. encouraging a fever, rather than suppressing it)*.[9, 12]

1.3.2 Principles

Traditional principles form the basis of naturopathic practice.[12, 13] These principles include: first, do no harm *(primum non nocere*), (supporting the) healing power of nature (*vis medicatrixnaturae)*, treat the cause (not just the symptoms) (*tollecausam)*, treat the whole person (rather than individual disease) (*tolletotum)*, doctor as teacher (to educate the patient) *(docere)*, disease prevention and health promotion, and wellness or wellbeing.[9, 13]

1.3.3 Theories

The theories are concepts which have been incorporated in the principles of naturopathic practice (e.g. treat the whole person) or which are used to guide naturopathic practice (e.g. value of a fever).[9] According to the World Naturopathic Federation, key theories that underlie naturopathic practice include:[9]

(1) Vital Force and Theory of Vitality – synonymous with the naturopathic philosophy of vitalism.

(2) Integration of the Individual – aligns with the naturopathic principle of treating the whole person.

(3) Naturopathic Cures – refers to the therapeutic concept of detoxification (e.g. fasting), revitalisation (e.g. in the form of mental therapy such as yoga), stabilisation (of an individual's health (e.g. through lymphatic drainage), and regeneration (e.g. in the form of mental therapy such as counselling).

(4) Value of a Fever – based on the understanding that fever helps the body fight an infection, and therefore, helping the body to heal itself.

(5) Therapeutic Order – refers to recommendation that naturopathic treatment is best applied in a certain order to resolve a patient's symptoms and address with the least potential for damage.

(6) Naturopathic Triad of Health – represented in the principle of 'treating the whole person' by addressing mind, body and spirit.

(7) Unity of Disease – all disorders can be traced back to three primary manifestations, namely: lowered vitality, abnormal composition of blood and lymph, and accumulation of waste materials, morbid matter and poisons.

(8) Hering's Law of Cure – stipulates the direction in which symptoms are cured: from the inside out, from the head down, from most important to least important organs, in reverse order of how they first appeared.

(9) Theory of Toxaemia – the main cause of disease is the accumulation of toxins (harmful materials or chemicals) from (e.g. too much stress or eating too much of the wrong foods).

(10) Emunctory Theory – elimination of toxins from the body is vital to achieve optimal health.

(11) Humoral Theory – spans all aspects of the naturopathic therapeutic encounter, including assessment, diagnosis, and treatment.

These philosophies, principles and theories focus on treatment and prevention of conditions, and promotion of health through naturopathic treatment modalities.

1.3.4 Modalities

In Australia, the most commonly prescribed modalities in naturopathic clinical practice include: nutritional medicine (e.g. nutraceuticals and supplements), dietary and lifestyle counselling and herbal medicine prescription.[12, 13] Some naturopaths also use homeopathy and manual therapies

(e.g. massage) as part of their practice. Naturopaths also report prescribing other interventions; such as meditation, yoga and exercise to support their patients.[1,9]

In a recent survey[1], Australian naturopathic practitioners reported that the most common modalities they use in their interventions are:

- lifestyle modifications (98% of practitioners)
- dietary modifications (90%)
- herbal medicine (90%)
- meditation (88%)
- exercise prescription (83%)
- yoga (75%)
- nutritional supplementation (65%) and
- homeopathy (36%).

1.4 HOW THE INTERVENTION MIGHT WORK

Naturopathic treatment often uses multi-modal interventions such as herbal medicine, nutritional supplementation, diet, and lifestyle modifications in combination with other supporting modalities, for example, homeopathy and manual therapies.[14] Some research suggests that the aforementioned whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice can improve health outcomes and improve quality of life in patients with chronic conditions or who are at risk of chronic conditions such as cardiovascular disease, chronic pain, type 2 diabetes and/or anxiety.[2, 15] Some interventions delivered or prescribed by naturopaths aim to improve patient's diet or lifestyle (e.g. exercise prescription, reducing intake of sugary or processed foods), with the health benefits of physical activity and a healthy diet well-documented in the scientific literature.[16]

According to the Australian Burden of Disease Study in 2015, 7.3% of the total burden of disease was due to poor diet while physical inactivity contributed to 2.5% of the total burden.[17] Using cardiovascular disease as an example, dietary risks contribute 40.2% of the total burden of disease, while alcohol use contributed 3.6%, tobacco use 11.5% and physical inactivity 8.0%.[17] However, the risk factors contributing to the burden of disease are not additive and have been described as having a 'joint effect', in view of the complex interactions between them.[17] Australian naturopaths may apply dietary advice and help develop a healthier diet based on the evidence-based Australian Dietary Guidelines[16] to improve a patient's risk of cardiovascular disease or other chronic conditions,[16-18] among counselling for other lifestyle modalities. Given the synergistic effect of smoking, poor diet and physical inactivity on chronic conditions,[19] adhering to lifestyle advice for these modalities may therefore have a synergistic effect on improving health.

The way naturopathy is practised may also enhance the effects of the naturopathic modalities administered or prescribed by the practitioner. The benefits of naturopathic practice may also arise from the practitioner-patient relationship.[14] For example, compared to family physicians, naturopaths practise with relatively longer consultation times with their clients.[20] This may enhance communication which in turn enhances adherence to therapeutic advice, including advice on lifestyle factors, although additional consultation time alone does not directly result in improved care.[21] However, longer consultation times may allow a naturopath to assess more of a patient's issues than a family physician is able to in a shorter consultation, which would influence patient-practitioner interactions.[20]

Although Zolnierek (2009) did not investigate naturopathic practitioners, their meta-analysis reported good physician communication is associated with greater patient adherence to treatment.[22] The rationale is that open communication and shared beliefs elicit clinical and psychosocial information from clients. Good communication also enables client involvement in decision-making and the discussion of benefits, risks, and barriers to treatment adherence, and develops rapport, trust and encouragement with clients.[14, 22] To further illustrate the relationship

between communication and treatment adherence, clients of complementary and alternative medicine (CAM) practitioners (which also encompass naturopathic practitioners) in Australia reported elements that helped change their health behaviour included the practitioner teaching them what to do, monitoring their progress, providing encouragement and directing them to information and resources they could use independently[23] All of these are components of good clinician communication.[22] The most frequently reported health behaviour changes made by clients of CAM practitioners in Australia were the lifestyle changes of improved diet and increased exercise.[23] As stated in Section 1.3, dietary advice was also the most common treatment provided by naturopathic practitioners, followed by Western herbal medicine, lifestyle advice and exercise advice.[24] Thus, where there is adherence to behavioural change advice (dietary, lifestyle and exercise advice), there may be resulting health benefits.

1.5 WHY IS IT IMPORTANT TO DO THIS REVIEW?

Australia has one of the highest rates of CAM practitioner use among developed countries and naturopathy is one of the most popular forms of CAM.[12, 25] The number of naturopathic consultations exceeds 4.9 million annually.[12] However, naturopathic practice is often accused of lacking an evidence base.[12, 26] Also, naturopathy is not regulated (i.e. it is self-regulated) which means any individual can currently practice naturopathy with or without appropriate training.[13, 27] Further research into effects of naturopathic practice and regulations, as practised in Australia, is required.[13] Hence, this Review will identify and evaluate the evidence for the clinical effectiveness of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice. The Review will inform the Australian Government's decisions about subsidies for natural therapies through private health insurance rebates.

The Australasian Cochrane Centre, Monash University was commissioned by the NHMRC to conduct an overview of systematic reviews to synthesise the effectiveness of naturopathy as a health service.[28] The overview, which was finalised in 2015, was part of the 'Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies'. It considered systematic reviews published between 2008 and May 2013. It identified one unpublished systematic review of whole system naturopathic medicine in chronic conditions. Of the 13 studies included in the unpublished systematic review, six were randomised controlled trials (RCTs), which were further assessed. These studies evaluated the effectiveness of naturopathic practice in cardiovascular disease, multiple sclerosis, anxiety and musculoskeletal pain. The primary outcomes included measures of pain, quality of life, anxiety and cardiovascular risk. The quality of the evidence was assessed as very low and the authors noted among the limitations that all of the studies were conducted in North America. Further, the authors concluded that while there was some evidence to suggest naturopathy as a health service improved patient health for a number of chronic health conditions, they urged caution in view of the differences in naturopathic practice, training and accreditation between North America and Australia.[28]

The 2015 overview did not include individual modal therapies used in naturopathic practice.[28] The Chair of the Department of Health's Natural Therapies Review Advisory Committee[29] noted that the authors may have missed systematic reviews that were published as grey literature, as searching was restricted to bibliographic databases. While the 2015 overview did not apply language restrictions in its search, its inclusion criteria limited studies to the English language only. The unpublished systematic review it identified had itself restricted languages to English, Spanish and French, in view of its North American focus.

This Review aims to evaluate and synthesise the evidence for the effectiveness of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice. In contrast with the 2015 overview, this Review will evaluate RCTs and non-randomised studies of interventions (NRSIs), and will include the most common (single) modalities of a therapy administered in the context of naturopathic practice.

1.6 OBJECTIVES

The main objective of the Review is to assess the effectiveness of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice for preventing, managing, treating and/ or delaying progression of health conditions in people with a clinical condition, pre-clinical condition or at risk of illness or injury.

2 Methods

This Chapter presents the methodology for conducting the Review. It describes the criteria for considering which studies are eligible, the search methods for identifying studies and the methods for collecting and analysing the data.

2.1 CRITERIA FOR CONSIDERING STUDIES IN THIS REVIEW

This section defines the types of studies that are considered eligible for inclusion in this Review. It describes inclusion and exclusion criteria of participants' characteristics, the study designs, interventions, comparators and outcome measures.

2.1.1 Types of studies

To be eligible, studies must be RCTs or NRSIs that examine the effectiveness of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice (the intervention). The intervention group must be able to be compared to a contemporaneous control or another intervention group. There is no minimum sample size for studies.

RCTs, as the main study type of interest, will include:[30]

- **Parallel RCTs of individuals:** where individual participants are randomly assigned to either an intervention group or to a control group
- **Cluster randomised trials:** where clusters of individuals, rather than individuals themselves, are randomised to different arms of the trial
- **Cross-over trials:** where the participants, upon completion of the course of one treatment are switched to another treatment, and
- **Pseudorandomised controlled trials (quasi-randomised controlled trials):** where participants are allocated to either the intervention or control groups by investigators using an allocation process that attempts randomisation, but where the method of randomisation is inadequate or not truly random (e.g. alternate allocation). This includes where the method of randomisation is not specifically stated or not strictly random. The methods of randomisation will be considered in the risk of bias assessment.

For cross-over trials, the concern as to whether wash-out periods are sufficient is acknowledged. Cross-over trials will be included where it is appropriate to do so in the context of the trial PICO, in line with Cochrane handbook Chapter 23.2.2: for example, where naturopathy interventions have a temporary effect and therefore are not likely to have carry over effects after the 'wash-out' period, which are used in the treatment of stable, chronic conditions. If cross-over trials are incorporated into a meta-analysis, they will be analysed separately to other RCTs (see section 2.3.9).

Where a cross-over trial is deemed to be an inappropriate design, only data from the first trial period will be included in the review. Unsuitable cross-over trials include:[31]

- where the medical condition evolves over time, such as a degenerative disorder, a temporary condition that will resolve within the time frame of the trial, or a cyclic disorder;
- when an intervention (or its cessation) can lead to permanent or long-term modification;
- if the elimination half-life of a drug is very long so that a 'carry-over' effect is likely; and
- if wash-out itself induces a withdrawal or rebound effect in the second period.

Specific NRSI study designs are eligible to ensure evidence can be evaluated for a broad range of populations and outcomes. For example, where RCTs are not available for a particular population, or where available RCTs only provide findings of low certainty.

NRSIs are also eligible where they include the following features:

- Data are collected from an 'intervention' group and a comparison group (e.g. control / placebo / other intervention) contemporaneously.
- Allocation (to 'intervention', control/placebo or another intervention) occurs by methods that are not random, including choice, availability, or non-random chance.

Consequently, the following types of NRSIs will be eligible (assuming they have the two features listed above) are:

- **Non-randomised controlled studies:** Experimental studies in which people are allocated to whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice or a control/placebo group and the outcomes compared.
- **Prospective cohort studies with contemporaneous comparator group:** Studies in which outcomes from a defined group of people (the cohort) are followed over time, with participants recruited before any intervention occurs. Outcomes are collected for both participants who are and are not exposed to whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice. The association between exposure and non-exposure with outcomes is examined.
- **Retrospective cohort studies with contemporaneous comparator group:** Studies in which outcomes from a defined group of people (the cohort) are examined after the intervention and outcomes occur. Outcomes are collected for both participants who are and are not exposed to whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice. The association between exposure and non-exposure with outcomes is examined.
- **Interrupted time series studies:** Studies that measure outcomes at multiple time points before and after an intervention (the 'interruption') is introduced. Outcomes from the intervention group are compared to those from the control group at the same time points. The design attempts to detect whether the intervention has had an effect significantly greater than any underlying trend over time.
- **Controlled before-and-after studies:** Studies that measure outcomes in an intervention group and control group before and after the implementation of an intervention. Outcomes for the intervention and control groups at the same time point are compared.
- **Case-control studies:** Studies that compare people with a specific outcome of interest ('cases') with people from the same source population but without that outcome ('controls'). The association between exposure and non-exposure to whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice with outcomes is examined.

The following types of studies are excluded:

- **Case series studies:** an uncontrolled observational study involving an intervention and outcome for more than one person.[30]
- **Case reports:** an uncontrolled observational study involving an intervention and outcome for a single person (or other unit).[30]
- **Cross-sectional studies:** studies that examine the relationship between diseases (or other health related characteristics) and other variables of interest as they exist in a defined population at one particular time. The temporal sequence of cause and effect cannot necessarily be determined in a cross-sectional study.[30]
- **Qualitative studies:** a research study that uses a qualitative method of data collection and analysis,[32] and
- **Single arm studies:** a sample of individuals is given the intervention and then followed over time to observe their response.[33]

Case reports, case series and single arm studies do not include contemporaneous comparator groups and do not provide evidence of effectiveness of the intervention. Cross-sectional studies allow for assessment of an association between intervention and outcome but do not provide evidence of cause and effect. Qualitative studies do not quantify effectiveness of the intervention which is required to conduct meta-analyses.

The study type of a publication will be confirmed by full text review by assessing study design features (Section 24.1 of the Cochrane Handbook[34]). It is acknowledged that the study design/type stated by the authors may not accurately reflect the actual study features: for studies of effectiveness, caution is required when assessing NRSIs according to existing evidence hierarchies, as the study labels from such hierarchies were originally derived from aetiological research questions and may not be applicable to the broad range of effectiveness studies (Section 24.2.1 of the Cochrane Handbook[34]).

Publication date

There will be no limitation to the publication date when the electronic searches for the systematic review are conducted. Studies provided to HealthConsult and Office of the National Health and Medical Research Council (ONHMRC) by Natural Therapies Review Expert Advisory Panel (NTREAP), NTWC or other stakeholders will only be excluded based on publication date where they are published after the search date of the electronic searches for each systematic review literature search. If they are published after the search date but are otherwise eligible, they will be included in the 'Characteristics of studies awaiting classification' table (see Appendix C, Table 8).

Trials that are ongoing, pending publication, or are completed but do not have results available, but otherwise meet inclusion criteria, will be listed as 'ongoing studies' and be documented in a 'Studies awaiting classification' table (see draft template at Table 8, Appendix C). Such trials may be published as conference abstracts.

Studies published in languages other than English

Databases in languages other than English will not be searched. However, studies published in languages other than English are not excluded from the Review. Refer to Section 2.3.1 for further details on how studies published in languages other than English will be screened.

2.1.2 Types of participants

Study populations may have any injury, disease, medical condition, or pre-clinical condition and can be any age. Studies with healthy but at risk populations are also eligible: this includes populations at risk of becoming ill or injured based on social risk factors (e.g. unemployment), biomedical risk factors (e.g. blood pressure), and behavioural risk factors (e.g. high alcohol consumption). There are no restrictions for study setting.

Studies that only include healthy populations seeking health improvement are excluded. However, where a study includes both healthy populations (ineligible) and eligible populations and separate data is available for eligible populations, the study will be included.

Searches will be limited to human studies, thus excluding animal and *in vitro* studies.

2.1.3 Types of interventions

Studies which meet the definition of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice are eligible. Both single and multi-modal interventions are eligible as long as they are described as naturopathy. To be eligible, the

naturopathy intervention must include at least one of the following modalities that are central to naturopathic care in Australia:

- herbal medicine,
- complementary medicine prescription,
- dietary advice or
- lifestyle advice.

There are no restrictions on the setting in which the naturopathy intervention is delivered, for example in person, by telehealth or by another medium. However, the intervention must be delivered in the context of naturopathic care, meaning at least one of the following criteria must be met:

(1) the study states the intervention was delivered by a naturopath
(2) the study states the intervention was delivered in the context of naturopathic care
(3) the intervention is described by the study as being a naturopathic intervention or by a naturopath.

Both single and multi-modal interventions that meet the definition of 'naturopathy' will be synthesised together, with subgroup analyses conducted if there is sufficient evidence within a population to do so.

Naturopathic interventions delivered as an adjunct to conventional care are eligible only if the comparator group also receives conventional care.

Whole system, multi-modal naturopathy interventions

Included are naturopathic interventions in which multiple modal interventions are delivered, including at least one modality considered central to naturopathic care in Australia (as above). Additional modalities may include (but are not limited to) yoga, meditation, exercise prescription, homeopathy and manual therapies such as massage, shiatsu and kinesiology.

Whole system, single modal naturopathy interventions

Included are naturopathic interventions in which one of the four modalities considered central to naturopathic care in Australia is delivered (dietary advice, lifestyle advice, herbal medicine, or complementary medicine).

Exclusions

Studies will be excluded where the ‘naturopathic practice’ intervention is combined with one or more other co-interventions, unless the effect of the naturopathic practice alone can be determined.

Studies are excluded if the intervention is a whole system, multi-modal therapy that does *not* meet the definition of naturopathic practice (for example, other traditional medical systems such as Traditional Chinese Medicine and Ayurveda). However, naturopathy interventions that include modalities derived from these systems (for example, yoga, acupuncture) will be eligible for inclusion if the intervention meets the definition of whole system, multi-modal ‘naturopathy’ as outlined above (i.e. includes at least one of the modalities central to naturopathic care in Australia).

Modalities that are not central to naturopathic practice in Australia (for example, yoga, meditation, exercise prescription, homeopathy and manual therapies such as massage, shiatsu and kinesiology), are excluded as single modal naturopathy interventions. However, these interventions are included when incorporated within a whole system, multi-modal naturopathy intervention that includes one or more interventions central to naturopathic care in Australia (see section 2.1.3 ‘Whole system, multi-modal naturopathy interventions’).

2.1.4 Types of comparators

The types of comparators used in studies will not be restricted. Placebo/ sham (if relevant), inactive control (i.e. inclusive of no intervention, wait list or usual care) and active comparators (i.e. inclusive of usual care or control if considered active) are eligible for inclusion. Analysis will be stratified by type of comparator:

(1) Placebo/sham (if relevant)
(2) Inactive Control (i.e. inclusive of no intervention, wait list or usual care)
(3) Active comparators, inclusive of usual care or control if considered active. Synthesis of evidence derived from studies with active comparators will be on the advice of the NTWC.

Where naturopathy is administered as an adjunct (for example, naturopathy plus standard care vs standard care alone) this will be classified as an inactive comparator and included in comparison (2) above.

The Review aims is to assess the effectiveness of naturopathic practice and not different types of naturopathic practice. Thus, studies that compare groups receiving one form of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice with groups receiving another type of naturopathic practice (analogous to a head-to-head trial) are not eligible for inclusion. This includes studies comparing multi-modal naturopathic care to single modal naturopathic care, two or more groups each being treated by different types of single modal naturopathic care, or studies that compare two or more groups each treated with different combinations of modalities in their multi-modal care.

Comparisons

Studies will be compared if they compare naturopathic care (the intervention) with, (1) placebo/sham (if relevant), an (2) inactive control as described above, with similar population groups and outcomes (see section 2.3.11). Subgroup analyses will be conducted for studies with multi-modal interventions and single-modal interventions. Studies with active comparators will be compared on the advice of the NTWC (see above (3)).

2.1.5 Types of outcome measures

Role of outcomes

Study eligibility will not be restricted by the type of outcomes measured and outcomes will not be used as an eligibility criterion. All included studies (regardless of their outcome measures or the time points of those outcome measures) will be included in the 'Characteristics of included studies table' (see Appendix C, Table 9). However, only certain outcomes will be extracted and examined in the analysis, as indicated by the NTWC prioritisation process.

Outcome domains

Due to the broad nature of the review, it is not possible to pre-specify outcomes for prioritisation in the review. To prioritise outcomes for data extraction and synthesis, NTWC will undertake a blinded prioritisation exercise.

Following the completion of screening, the NTWC will be provided with a list of conditions, outcome domains and outcome measures to prioritise. The list will be based on outcomes reported in the included studies as well as outcomes included within a relevant core outcome sets (if available), identified using Core Outcome Measures in Effectiveness Trials (COMET)[35] and in other relevant or related Cochrane reviews. No additional information will be provided that could enable NTWC to identify the included studies (e.g. names of authors, the study the country was conducted in), the results of the studies, or the number of studies examining each condition, outcome domain or outcome measure.

Throughout the outcome prioritisation exercise, the NTWC will apply Grading of Recommendations Assessment, Development and Evaluation (GRADE)[36] principles to identify up to seven critical and/or important (but not critical) outcomes for each population condition, for HealthConsult reviewers to extract and report on.

Outcome measures and timepoints of interest

The NTWC will focus on the relevance and validity of outcome measures. As the Review is focused on assessing the clinical effectiveness of therapies, the prioritised outcomes will relate to the potential benefits of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice rather than potential harms. As stated under 'Outcome domains', the NTWC's approach to outcome prioritisation will be blinded. Outcome measures based on personal health care preferences, patient satisfaction, safety, quality, and economic outcomes (e.g. cost-effectiveness) are out of scope, as are adverse effects of treatment. However, effectiveness outcomes that show a harm through worsening of symptoms are eligible for inclusion. Timepoints will be pre specified during the outcome prioritisation exercise conducted by the NTWC. Ideally, the chosen timepoint will be a clinically important time point (see also Section 2.3.9, Repeated observations).

2.2 SEARCH METHODS FOR IDENTIFICATION OF STUDIES

2.2.1 Electronic searches

The proposed search strategies have been adapted from Myers (2019)[2] and Cooley (2012).[37] Search strategies based on whole system naturopathy (including single and multi-modality interventions) were developed for each database and include search terms for the study types RCTs and NRSIs with contemporaneous control groups. The searches have been designed so that they are restricted to humans but are not restricted by population, outcome, date, language or geography.

Search strategies have been designed for the following databases:

- Medline via OVID
- Embase
- Cochrane CENTRAL
- Allied and Complementary Medicine Database (AMED)
- Cumulative Index of Nursing and Allied Health Literature (CINAHL)

Databases in languages other than English will not be searched.

The proposed search strategies for each database are included in Appendix B. Database-specific terms for Medline, Embase, AMED, CINAHL and Cochrane CENTRAL are included. This included the use of a controlled vocabulary (refer to Section 4.4.4 of the Cochrane Handbook[38]) to map search terms to the individual database's subject headings, as MeSH (Medline and Cochrane) terms. The subject headings used in other databases such as Embase are not identical. This was undertaken by entering (for example) a MeSH term into a database and then selecting the appropriate subject headings the database maps the term to. Changes to search terms for text words or all field codes were not required.

2.2.2 Other searches

The reference lists of included articles will be checked to find additional eligible studies (backwards citation search).

Studies provided by the public and key stakeholders (via the Department of Health), the NTREAP and the NTWC will also be screened for eligibility. Where these groups recommend particular systematic reviews, they will be examined for eligible RCTs and NRSIs. The ONHMRC will also provide the 2015 evidence evaluation for whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice to HealthConsult reviewers to identify eligible primary studies within included systematic reviews. Systematic reviews not published in English will not be translated, but will be examined to identify eligible RCTs and NRSIs.

The following clinical trials registries will be searched: International Trials Registry; Australian New Zealand Clinical Trials Registry; and ClinicalTrials.gov.

Studies and publications that have been identified for inclusion by full text review will be checked for retraction or errata within the databases. For example, in Medline using the search strategy 'retracted publication.pt. or retraction of publication.pt.' (refer to Section 3.9 of 4.S1 Technical Supplement to Chapter 4 of the Cochrane Handbook[39]) together with the citations for the eligible studies. Appropriate terms for Embase, the Cochrane library, CINAHL and AMED will be included.[39]

2.3 DATA COLLECTION AND ANALYSIS

2.3.1 Selection of studies

Results from electronic database searches

Screening of citations will be conducted using Endnote software. Following the database searches, duplicate citations will be removed. The studies will then be screened for eligibility by title and abstract and then by full text.

At the title and abstract stage, studies will be screened for eligibility by a primary reviewer, with a secondary reviewer independently assessing an initial 20% of citations.[40] The duplicate screening aims to achieve 80% inter annotator agreement between reviewers and will progress until this is achieved. Where there are discrepancies that the two reviewers cannot agree on, a third reviewer's opinion will be sought.

At the full text stage, two independent reviewers will screen the reports for eligibility: any disagreements will be discussed between the primary and secondary reviewer, with a third reviewer consulted should the former be unable to reach an agreement.

Eligibility criteria

If multiple reports of the same study are identified, they will be collated so that the study rather than each report is the unit of interest in the Review.[38] It is anticipated that multiple reports of a unique study will be identified by their having the same trial identification number, similar author names (reflecting the same research team of a study), locations and settings, intervention details, the numbers of participants and their baseline data and the date and duration of the study.[38] This will ensure that data from a study is included only once in the analyses and syntheses of evidence.

For citations of studies published in languages other than English that may be eligible for inclusion according to screening at the title and abstract stage (see below), full text publications will not be translated but will be documented in a 'Studies awaiting classification' table (see draft template at Table 8, Appendix C).

Where it is unclear whether a study meets the meet the definition of whole system, multi-modal or single modal interventions delivered in the context of naturopathic practice, HealthConsult will seek further guidance from the NTWC (via NHMRC). NTWC will be presented with the minimum

information required to enable them to determine whether the intervention meets the definition. NTWC will be blinded to identifying details of studies (e.g. study type, sample size, findings).

A PRISMA flowchart will be generated to document the results of the searching and screening process. Studies excluded at the full text stage will be tabulated with a corresponding rationale for exclusion.

Citations provided by NTWC, NTREAP and stakeholders

The number of primary RCTs and NRSIs and the number of systematic reviews (and the number of primary studies they include) received from NTWC, NTREAP and other stakeholders will be reported in the Review. The number of unique eligible primary studies (that is, studies not identified in the electronic database searches) will be reported.

If systematic reviews are provided by NTWC, NTREAP or other stakeholders, the primary RCTs and NRSIs they include will be cross-checked against the citations retrieved in the electronic searches. Any RCTs or NRSIs, including those provided by NTWC, NTREAP or other stakeholders, that were not identified in the electronic searches will be screened by title and abstract and then by full text as described above under *Results from electronic database searches*.

Citations for primary RCTs or NRSIs and 'grey literature' provided by NTWC, NTREAP or other stakeholders will be screened for eligibility as full text citations. As above, two independent reviewers will screen the full text reports for eligibility. Any disagreements will be discussed between the primary and secondary reviewer, with a third reviewer consulted should the former be unable to reach an agreement.

Studies published in languages other than English

Studies published in languages other than English that are retrieved will be managed using the following protocol:

(1) Database searches will not be restricted by language.
(2) If the title and abstract are not available in English, they will be translated using Google translator or an equivalent method (then proceed to step 4).
(3) If online translation does not facilitate an understanding of the title and abstract, the citations for these studies will be listed as 'studies unable to be translated or interpreted at the title/abstract stage'.
(4) Translated titles and abstracts will be screened and citations that are not eligible excluded. The number of these excluded citations published in languages other than English will be reported in the results of the search and in the PRISMA chart.
(5) If the translated citation indicates the study is likely to meet the criteria for considering studies for inclusion in the review (based on title and abstract screen), or if there is any uncertainty:
    - The translated citation and available information will be recorded in the 'Studies awaiting classification' table to inform readers of the Review of the availability of other possibly relevant reports. The information will also be reflected in the PRISMA flow diagram.
    - A copy of the finalised 'Studies awaiting classification' tables (see Table 8, Appendix C) will be provided to ONHMRC, noting that the review is not expected to include any of these articles.
    - The potential risk of language bias and the implications in the Evidence Evaluation Report will be appraised.
    - Appropriate qualifying statements will be presented throughout the Evidence Evaluation Report that acknowledge only the evidence published in English was reviewed.
    - Potential limitations due to language bias and the potential impact on certainty of evidence will be presented in the Evidence Evaluation Report, noting that they may influence the conclusions of the Review.

2.3.2 Data collection process

Information, including results data will be extracted into tables (Word software) by two reviewers independently (see data extraction templates in Appendix C, Table 9 and Table 10). Disagreements will be resolved by discussion between the two reviewers, with a third reviewer consulted if agreement cannot be reached between the former.

The table templates will first be pilot tested by the two reviewers responsible for data extraction (Table 9 and Table 10). Each reviewer will independently extract a sample of three representative primary studies which will represent the range of study types and PICO eligible for inclusion within the review. The completed tables will be checked by the lead reviewer and any necessary changes to the templates or the instructions for using them will be made. This will allow for consistency in the data extraction process across reviewers.

2.3.3 Requests for data

Where feasible (i.e. authors contact details are included in the publication) and if relevant, any missing information from the included studies will be sought from the original authors. Two contact attempts by email will be made to the author prior to stating the data as missing in the evidence evaluation and/or technical report.

2.3.4 Data items

The data to be extracted from eligible studies are:[41]

- Study citation
- Year of publication
- Study type: RCTs or NRSI with contemporaneous comparator (e.g. cohort or case-control studies)
- Study duration
- Country
- Population group, number of participants, setting, and demographic data (including gender, age, condition and/or diagnosis), inclusion and exclusion criteria
- Intervention:[41, 42]
    1. Name and description of intervention
    2. Description of rationale, theory or goal of the elements essential to the intervention
    3. Materials used in the intervention
    4. Procedures used in the intervention
    5. Intervention provider
    6. Modes of delivering the intervention
    7. Location where the intervention occurred
    8. Timepoints the intervention was delivered, time period, frequency/number of sessions, duration of intervention session, intensity, dosage
    9. Tailoring of the intervention, if the intervention was personalised, titrated or adapted, the rationale and method for doing so
    10. Modifications to intervention, when they occurred, why and how
    11. Strategies to maintain or improve adherence/fidelity to intervention, if assessed
    12. Actual adherence or fidelity to intervention, if assessed.
- Comparator:[41, 42]
    1. Name and description of comparator
    2. Description of rationale, theory or goal of the elements essential to the comparator
    3. Materials used in the comparator
    4. Procedures used in the comparator
    5. Comparator provider
    6. Modes of delivering the comparator

7. Location where the comparator was administered
8. Timepoints the comparator was delivered, time period, frequency/number of sessions, duration of comparator session, intensity, dosage
9. Tailoring of the comparator, if the comparator was personalised, titrated or adapted, the rationale and method for doing so
10. Modifications to comparator, when they occurred, why and how
11. Strategies to maintain or improve adherence/fidelity to comparator, if assessed.
12. Actual adherence or fidelity to comparator, if assessed.

- Outcomes: The outcome results may be continuous or dichotomous (categorical) and will include 'critical' and 'important but not critical' outcomes (up to seven) and the timepoints at which they were measured.
  - Outcomes:
    1. Primary outcomes: Description, including measurement method
    2. Secondary outcomes: Description, including measurement method
    3. Whether there is evidence that the outcome domain was assessed (especially important if the outcome was assessed but the results not presented)
    4. Measurement tool or instrument (including definition of clinical outcomes or endpoints); for a scale, name of the scale, upper and lower limits, and whether a high or low score is favourable, definitions of any thresholds if appropriate
    5. Specific metric (e.g. post-intervention anxiety, or change in anxiety from baseline to a post-intervention time point, or post-intervention presence of anxiety (yes/no))
    6. Method of aggregation (e.g. mean and standard deviation of anxiety scores in each group, or proportion of people with anxiety)
    7. Timing/timepoints of outcome measurements (e.g. assessments at end of eight-week intervention period, events occurring during the eight-week intervention period)
    8. For each group, and for each outcome at each time point: number of participants randomly assigned and included in the analysis; and number of participants who withdrew, were lost to follow-up or were excluded (with reasons for each)If subgroup analysis is planned, the same information would need to be extracted for each participant subgroup
  - Results data:
    1. Intervention results: participant number, mean/proportion
    2. Comparator (point estimate): participant number, mean/proportion
    3. Point estimate: risk estimates and direction of effect. This will included estimates that are adjusted for confounders, if reported by the study.
    4. Measures of variation: such as standard deviation, standard error and 95% confidence intervals.
       - Other: funding sources, notable conflicts of interest of trial authors.

Refer to Table 9 and Table 10 in Appendix C for information and data extraction templates.

Data required to assess studies' risk of bias will also be extracted (see sections 2.3.6 and 2.3.7 and Table 11 to Table 13 in Appendix C).

2.3.5 Missing data

Data will be analysed on an intention to treat (ITT) basis where possible, including when conducting meta-analyses. Missing data such as standard deviations will be imputed if the study reports sufficient results to permit confidence in such calculations, for example using the calculation methods in sections 6.5.2.2 and 6.5.2.3 of the Cochrane Handbook[43]. Where possible and relevant, the study authors will be contacted to obtain missing data, as described in Section 2.3.3. If assumptions are made regarding any of the methods used to address missing data (for example, that the data are assumed missing at random), they will be stated clearly. Sensitivity analyses (refer to Section 2.3.14) will be performed to assess how sensitive the results are to reasonable changes in the assumptions made. Risk of bias due to missing outcome data will be

assessed as per Sections 2.3.6, 2.3.7 and 2.3.13. The potential impact of missing data on the findings of this Review will also be addressed in the Discussion section (section 10.12, Cochrane Handbook[44]).

2.3.6 Tools to address risk of bias in individual studies

Randomised controlled trials and pseudorandomised controlled trials

The risk of bias will be assessed using the Cochrane Risk of Bias Tool 2 (ROB2) for RCTs.[45, 46] This tool is comprised of five domains and if a study assesses more than one outcome, the tool is applied to each relevant outcome separately:

(1) Risk of bias arising from the randomisation process.
(2) Risk of bias due to deviations from the intended interventions (effect of assignment to intervention or effect of adhering to intervention).
(3) Risk of bias due to missing outcome data.
(4) Risk of bias in measurement of the outcome.
(5) Risk of bias in selection of the reported result.

The ROB2 provides an assessment template for the five domains with signalling questions in each domain.[47] For each domain, the tool provides an algorithm to rate whether the risk of bias is low, of some concern or high and an algorithm is also provided to rate the overall risk of bias as low, of some concern or high.[45] For each outcome, the risk of bias assessment will be reported in the template shown at Table 11 (see Appendix C). The judgement of the risk of bias for each of the five domains (high risk, some concerns, low risk) and a rationale for the judgement will be reported, in addition to the overall risk of bias for the outcome. The overall risk of bias for the outcome will also be reported in the outcomes data extraction template (see Table 10 in Appendix C).

The effect of interest will be assignment to the interventions at baseline (i.e. intention to treat).

Non-randomised studies of interventions

The risk of bias for NRSIs will be assessed using:

- the Risk Of Bias In Non-randomised Studies of Intervention (ROBINS-I) tool[48] for all NRSIs except for case-control studies
- the SIGN checklist for case-control studies version 2.0[49], for case-control studies.

The ROBINS-I tool assesses risk of bias for a particular outcome of a study in the following seven domains:

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

The ROBINS-I provides a template for assessing all seven domains with signalling questions for each domain.[50] For each domain, an algorithm determines whether the risk of bias is low, moderate, serious, critical or that there is no information on which to base a judgement for that domain and another algorithm determines the overall risk of bias for the study's outcome, as low, moderate, serious, critical, or that there is no information on which to base a judgement about risk of bias.[48] For each outcome in an NRSI (with the exception of case-control studies), the risk of bias assessment will be reported in the template shown at Table 12 (see Appendix C). The judgement of the risk of bias for each of the seven domains (low risk, moderate risk, serious risk,

critical risk or no information) and a rationale for the judgement will be reported, in addition to the overall risk of bias. The effect of interest will be assignment to the interventions at baseline.

The ROBINS-I tool requires that potential confounding domains and co-interventions be pre-specified. Potential confounding domains and co-interventions will be identified for each population/condition; where studies examine very different outcomes for the same population/condition, relevant confounding domains and co-interventions to these outcomes will also be identified. NTWC advice will be sought to identify relevant and typical confounding domains and co-interventions as required. In addition, the ROBINS-I tool requires that potential confounding domains and co-interventions specific to the study being assessed or which were identified as important by the study authors be reported.

If an NRSI is assessed as being at critical risk of bias in any one domain, its details will be recorded in characteristics of included studies table and the reason for critical risk of bias rating documented, but it will not be further assessed and will not contribute to data synthesis.

It is acknowledged that the ROBINS-I tool was developed to align more closely with NRSIs of cohort-like study designs.[48] While signalling questions are yet to be published for before-after studies (e.g. time-interrupted series studies), the Cochrane handbook provides guidance on applying the tool to these designs (Table 1, see Table 25.6a of the Cochrane Handbook[51]).

Table 1: Controlled before-after studies (including time-interrupted series) – considerations in ROBINS-I risk of bias assessments[51]

ROBINS-I risk of bias domain	Considerations in assessing bias domain
Confounding	Whether: • measurements of outcomes were made at sufficiently many time points, in both the intervention and comparator groups, to permit characterization of pre-intervention trends and patterns; • any extraneous events or changes in context around the time of the intervention that could have influenced the outcome were experienced equally by both intervention groups; and • pre-intervention trends and patterns in outcomes were analysed appropriately and found to be similar across the intervention and comparator groups.
Selection of participants into the study	The issues are similar to those for follow-up studies. For repeated cross-sectional surveys of a population, there is the potential for selection bias if changes in the types of participants/units included in repeated surveys differ between intervention and comparator groups.
Deviations from intended interventions	[If assessing on the basis of assignment to intervention] The issues are the same as for follow-up (cohort) studies.
Missing data	Whether outcome data were missing for whole clusters as well as for individual participants.
Measurement of the outcome	Whether: • methods of outcome assessment were comparable across intervention groups and before and after the intervention; and • there were changes in systematic errors in measurement of the outcome coincident with implementation of the intervention.
Selection of the reported result	The issues are the same as for follow-up studies.

Reproduced with permission, Cochrane Handbook version 6.1, 2020

There is at present little guidance on how to adapt the ROBINS-I tool for case-control studies, although it is recognised that recall bias contributes to the misclassification of intervention status.[51] It is thus proposed to assess the risk of bias in case-control studies using the SIGN checklist for case-control studies version 2.0.[49] The SIGN checklist assesses internal validity of case-control studies through 11 questions (see Table 13, Appendix C) that cover the domains of subject selection, assessment measurement, confounding and statistical analysis. The checklist then provides an overall assessment of how well the study minimised the risk of bias or confounding. Judgements for each of the items in the checklist will be reported for each outcome, as well as the

overall assessment. A case-control study will be recorded in the characteristics of included studies table but it will not be further assessed and will not contribute to data synthesis if:

- the overall assessment of how well the study minimised risk of bias or confounding is judged as unacceptable
- the study does not include clear definitions of the source population
- the study does not comment on how cases or controls were selected
- the outcome measures are not stated or the study bases its main conclusions on secondary outcomes; or the study does not address the possibility of confounding.

The reason for exclusion from data synthesis will be included in the 'Characteristics of included studies' table.

2.3.7 Risk of bias assessment process

The risk of bias will be assessed by two reviewers independently. Any disagreements will be discussed between the two reviewers, with a third reviewer consulted should the former be unable to reach an agreement.

The risk of bias assessment tools will be piloted for a sample of three RCTs (Cochrane Risk of Bias tool), three NRSIs including cohort, before-after and time-interrupted series studies (ROBINS-I tool) and three case-control studies (SIGN) to include the range of study types and PICO eligible for inclusion within the review. The two reviewers responsible for data extraction will complete the risk of bias assessments and the lead reviewer will check the completed assessments and discuss how differences can be resolved. This will allow for consistency in the risk of bias assessment process, including the management of differences in identifying potential confounders according to the ROBINS-I tool.

If there are sufficient studies, risk of bias will also be addressed by sensitivity analyses (refer to Section 2.3.14).

2.3.8 Measures of effect

Measurement of treatment effect will be by mean differences (MD) (preferred) or standardised mean differences (SMD) and 95% confidence intervals (continuous outcomes) and risk ratios (preferred) or odds ratios with 95% confidence intervals (dichotomous data) (refer to Sections 6.4 and 6.5 of the Cochrane Handbook 2019[43]). Hazard ratios will be extracted for studies that assess time to event (refer to Section 6.8, Cochrane Handbook 2019[43]). To reduce effects of confounding, summary statistics from NRSIs will be reported as adjusted effect estimates where available. The minimally clinically important difference (MCID) will be sourced from published reports where possible, or will be guided by advice from the NTWC; alternatively, a 25% relative risk increase of 25% or more will be used as a default threshold for appreciable harm or benefit (GRADE Handbook section 5.2.4.2).

For continuous outcomes, standardised mean difference (SMD) will be calculated, if not already reported by studies, prior to analysis and synthesis by meta-analysis (section 6.5.1.2 of the Cochrane Handbook[43]). This is appropriate where an outcome has been measured using different scales in different studies.

2.3.9 Unit of analysis issues

For eligible RCTs or NRSIs, the unit of analysis is the individual participant. Attention will be given to cross-over RCTs and cluster RCTs to avoid under- or -overestimating precision (refer to Section 6.2, Cochrane Handbook 2019[43]).

Cluster randomised controlled trials

If the study is a cluster RCT, the unit of analysis will be the cluster reported by the RCT. If they are included in a meta-analysis, then the cluster will be treated as though it is a single individual, using summary measurements from each cluster (Cochrane Handbook section 23.1.3[31]). It is acknowledged that if clusters vary in their size, the precision of the effect estimate may be reduced. Thus, if a meta-analysis includes cluster RCTs, sensitivity analyses will be performed by excluding the cluster RCTs (refer to Section 2.3.14).

For cluster RCTs where the report has inappropriately analysed its results as though it is the individual that is randomised (i.e. unit-of-analysis error), they will only be included in meta-analyses if they present sufficient information that allow for approximate correct analyses (Cochrane Handbook section 23.1.4[31]):

(1) the number of clusters (or groups) randomized to each intervention group and the total number of participants in the study; or the average (mean) size of each cluster;
(2) the outcome data ignoring the cluster design for the total number of individuals (e.g. the number or proportion of individuals with events, or means and standard deviations for continuous data); and
(3) an estimate of the intracluster (or intraclass) correlation coefficient (ICC).

An effective sample size can then be calculated to reduce the size of the cluster trial for meta-analysis: 1 + (average cluster size – 1) x ICC.

Cross-over trials

Cross-over trials will be included where it is appropriate to do so in the context of PICO criteria and in line with section 23.2.2 of the Cochrane Handbook[31] (see section 2.1.1). There are three approaches to incorporating cross-over trials in meta-analyses, each also with potential unit of analysis issues (Cochrane Handbook section 23.2.6):[31]

(1) treat all intervention measurements and all comparator measurements as though the trial were a parallel RCT. The Cochrane Handbook recommends avoiding this approach as unit of analysis errors may arise from the confidence intervals being too wide resulting in the trial receiving too little weight and clinically important heterogeneity may be disguised.
(2) include only the data from the first period prior to the cross-over (equivalent to a parallel RCT). This subset of data may be biased if its reporting by researchers depends on their having found statistically significant carry-over. However, it may be appropriate if carry-over is thought to be a problem or if the cross-over design is considered inappropriate (see section 2.1.1).
(3) impute missing standard deviations for inappropriately reported cross-over trials to approximate a paired analysis. However, the suitability of this method depends on the confidence in the imputed standard deviations and how robust the meta-analysis results are to plausible imputed results (section 23.2.7 of the Cochrane Handbook[31] and section 2.3.5 Missing data).

It is proposed to utilise approach (2): it is acknowledged there are potential difficulties and concerns with assessing cross-over RCTs and that even only including the first study period prior to the cross-over (the equivalent of a parallel RCT) may still pose an unacceptable risk of bias (Section 23.2.3, Cochrane Handbook 2019[31]). Thus, if there are a sufficient number of cross-over RCTs that meet the inclusion criteria, a meta-analysis will be conducted separate to meta-analyses of parallel RCTs (Section 23.2.8, Cochrane Handbook 2019[31]).

Repeated observations

For studies that assess the same outcome several times over a long duration, it is acknowledged that results from more than one timepoint for each study cannot be combined in a standard meta-analysis without a unit-of-analysis error (Section 6.2.4 of the Cochrane Handbook[43]). Timepoints will be pre-specified during outcome prioritisation exercise conducted by the NTWC. Ideally, the chosen timepoint will be a clinically important timepoint.[43] It is acknowledged that choosing a timepoint that maximises the data available may lead to reporting biases.[43]

Naturopath practitioners treating multiple trial participants

A unit-of-analysis error could arise in an individually randomised trial where naturopath practitioners each treat multiple participants. Outcomes of participants of the same naturopath practitioner will be correlated, and if this correlation is not accounted for in the study analysis, the study's standard error may be incorrect which has implications for the meta-analysis.

If such intervention-related clustering is identified, it will be described narratively with consideration given to any impacts of clustering in the Review findings. No imputation or statistical methods will be used to adjust for this clustering.

2.3.10 Studies with more than two intervention groups

For eligible studies with more than two groups, results from treatment arms that do not meet the criteria for the intervention or comparator will be excluded from the analyses (section 23.3.2, Cochrane Handbook 2019[31]). If all groups are eligible or more than two groups are eligible, they will be included in the analyses. Care will be taken where a study with more than two groups is included in a meta-analysis, so that the participants from the one treatment arm are only included once (section 23.3.4, Cochrane Handbook 2019[31]). Where appropriate and possible, groups will be combined to allow for pair-wise comparisons in a meta-analysis (section 23.3.4, Cochrane Handbook 2019[31]).

2.3.11 Meta-analysis

Both single and multi-modal interventions that meet the definition of 'naturopathy' will be synthesised together, with subgroup analyses conducted if there is sufficient evidence within a population to do so. See section 2.1.3 for definitions of whole system, multi-modal and whole system, single modal interventions.

The primary comparison of interest will be naturopathic care vs placebo/sham (if relevant). synthesis will be undertaken for studies that compare naturopathic care with either Placebo/ sham (if relevant) or (in the absence of studies measuring placebo/sham); inactive control (including no intervention/treatment, wait list or usual care.

Sensitivity analyses will be performed to investigate the robustness of the treatment effect by performing analysis that includes all trials combined, i.e. trials with placebo/sham or inactive control to see if inclusion of trials that did not blind participants changed the overall treatment effect.

Results data from studies comparing naturopathic care with 'other' interventions will be extracted and presented in data tables, but will not be synthesised further, except where requested by the NTWC. These data will be presented as an 'evidence inventory' to provide a snapshot of the available evidence comparing naturopathic care with 'other' interventions.

The NTWC may request that data comparing naturopathic care with an 'other' intervention be synthesised, where:

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

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

When reviewing full text articles, studies (including conference abstracts) with no useable results data will not be included in any syntheses: this may occur when data is not presented in a useable

form, for example, when study authors state there is no difference in an outcome but do not provide point estimates that can be assessed. Studies that state they collected data on an outcome but do not report on it will be considered under reporting bias (section 2.3.13). A hand search will be undertaken to locate eligible additional publications based on the same study that have not already been identified.

Assessing the evidence for heterogeneity

Where there are at least two RCTs or two NRSIs that are sufficiently similar in population and outcome that can be included in a meta-analysis and that meet the criteria for quantitative synthesis as described in Section 2.3.11 heterogeneity will be quantified and assessed using a chi-squared test with a significance level of p=0.1 and by I^2 statistics (Cochrane Handbook section 10.10.2[44], calculated by RevMan version 5.3).[52] Forest plots will also be generated and a visual inspection of the overlap of the confidence intervals will be taken into consideration when assessing heterogeneity.

As per Section 10.10.2 of the Cochrane Handbook,[44, 53] the I^2 value will be interpreted as follows:

- 0% to 40%: might not be important
- 30% to 60%: may represent moderate heterogeneity
- 50% to 90%: may represent substantial heterogeneity
- 75% to 100%: considerable heterogeneity.

A low p value in the chi-squared test indicates evidence of heterogeneity of intervention effects.[53]

Heterogeneity will be reported descriptively for other studies that are not included in a meta-analysis and therefore do not have an I^2 statistic calculated.

Synthesis of RCTs

Meta-analyses using RevMan[52] will be performed if there are at least two RCTs with similar outcome measures and population and which meet the criteria for quantitative synthesis as described in Section 2.3.11 The analyses will be conducted for studies with control (inactive) comparators. Studies with 'active' comparators will be analysed on advice from NTWC (see Section 2.1.4). See Section 2.3.8 regarding the treatment measurements to be extracted for analysis. Random effects (DerSimonian and Laird method, inverse variance)[52] models will be used to calculate confidence intervals and to incorporate the heterogeneity among studies (refer to Cochrane Handbook Sections 10.10.3 and 10.10.4[44]).

Synthesis of NRSIs

Meta-analyses using RevMan 5.3[52] will be performed if there are at least two NRSIs of similar outcome measures and population and which meet the criteria for quantitative synthesis as described in Section 2.3.11 . Studies assessed as having critical risk of bias according to ROBINS-I or for case-control studies assessed as rejected/unacceptable according to the SIGN checklist will be excluded from synthesis (see section 2.3.6). The analyses will be conducted for studies with control (inactive) comparators. Studies with 'active' comparators will be analysed on advice from NTWC (see Section 2.1.4). NRSIs will be included in a single meta-analysis for a particular comparison and outcome, with subgroup analyses by different NRSI types (e.g. cohort studies, case-control studies) and by overall risk of bias judgement. See Section 2.3.8 regarding the treatment measurements to be extracted for analysis. Random effects (DerSimonian and Laird method[52]) models will be used to incorporate heterogeneity among studies (refer to Cochrane Handbook sections 10.10.3 and 10.10.4[44]).

Methods for calculating heterogeneity variance are detailed above under 'Assessing the evidence for heterogeneity'.

2.3.12 Summary and synthesis when meta-analysis is not possible

Evidence from RCTs and NRSIs where a meta-analysis is not appropriate and/or possible (refer to 12.1 Cochrane handbook[54]) will be synthesised as described below.

Sufficiently-powered studies with low risk of bias will not be excluded from the synthesis but the synthesised result will comprise two parameters:

(1) Where there are studies that are sufficiently powered (i.e. have a proportionate sample size to the research question e.g. measuring the difference between groups), with a low risk of bias, which cannot be included in meta-analyses, their results will be emphasised in the narrative description and presented in the comments section of the Summary of Findings table. For example, the results may be reported as, 'Results from one study (N=550) at low risk of bias: RR 0.71 (0.27, 1.88), p-value (e.g. 0.05), favours intervention.'
(2) The results of the remaining studies (i.e. those not sufficiently powered) will be summarised using vote-counting (based on direction of effect, not statistical significance, refer to section 12.2.1.3, Cochrane Handbook[54]). For example, the results may be reported as, '3/5 of the remaining studies reported an effect in favour of the intervention'. Evidence of an effect will be assessed using the Sign Test.

If there are no sufficiently powered studies with low risk of bias, then vote-counting will be conducted for all studies (based on direction of effect). Vote-counting results will also be presented narratively in the comments section of the Summary of Findings table for the relevant outcome. Results will also be displayed as a forest plot without the summary diamond, to provide a visual display of each individual studies' summary statistics.

2.3.13 Risk of reporting bias across studies

The selective reporting or under-reporting of outcomes will be assessed for individual RCTs and NRSIs as part of the risk of bias assessment (refer to Sections 2.3.6 and 2.3.7).

The non-reporting of evidence refers to when evidence is not available. This includes studies which state in their methodology or trial protocol that they will assess certain outcomes, but then do not report the results of those outcomes or do not report the outcomes in a useable form for data synthesis (e.g., may only state there was no difference or that there was a significant difference but without presenting means or point estimates).

Studies that may exhibit non-reporting of evidence will be identified when published reports are screened at the full-text stage (section 2.3.1), where studies report no useable results for the outcomes prioritised by the NTWC, and also when trial registries are searched (section 2.2.2, Chapter 13.2 of the Cochrane Handbook[55]). This will assist in reducing bias due to non-reporting of evidence. Such studies will be documented in the data extraction table (see template Table 10, Appendix C) and that their outcome results are missing; if significance or direction of effect are reported, they will be included here.

In line with signalling questions 4.1, 4.2 and 4.4 of the preliminary Risk of Bias due to Missing Evidence tool (ROB-ME)[56] and with section 13.3 of the Cochrane Handbook[55], the impact of risk of bias attributed to non-reporting of results on data syntheses will be appraised by comparing studies that did not report results with studies that were included in a meta-analysis for a particular comparison and outcome according to:

- The number of studies;
- The sample sizes of the studies;
- Direction of effect and/or p value if reported.

From these comparisons, the Review will formulate a conclusion as to whether the non-reporting of results for a particular comparison and outcome contributes a substantial risk of bias. For example, if studies that do not report results for outcomes of interest have large sample sizes or

there are more of them than there are studies that can be incorporated into a meta-analysis, there is likely to be risk of bias in the data synthesis.

Non-reporting bias as publication bias will also be assessed via regular funnel plots (section 5.2.5, GRADE Handbook 2013[36]), where there are at least 10 RCTs (sections 13.3.5.2 and 13.3.5.5 of the Cochrane Handbook 2019[55]) with similar populations, interventions and outcome measures with which to conduct a meta-analysis. However, it is acknowledged that there are limitations with funnel plots as an index of publication bias and undue weight will not be placed on their findings (sections 13.3.5.2 and 13.3.5.5 of the Cochrane Handbook[55]).

It is acknowledged that small studies may be awarded relatively more weight in a random effects analysis (section 10.10.4 of the Cochrane Handbook[44]). The impact of such studies, if there are sufficient studies overall to conduct a meta-analysis, will be tested in sensitivity analyses (see section 2.3.14).

2.3.14 Addressing risk of bias

If there are sufficient eligible RCTs to conduct meta-analyses, risk of bias will be addressed by sensitivity analyses (refer to section 7.6.1 of the Cochrane Handbook[57] and discussed below). Where there are outcome measures and populations with less than two RCTs and/or NRSIs, the studies will be presented with a narrative discussion of the risk of bias across the studies.[57] The GRADE system (refer to Section 2.3.16) will be used as an explicit measure of the certainty of evidence across the studies to help ensure that judgements about the risk of bias are considered appropriately when interpreting the results of the review.[57] This will also help ensure other factors affecting the quality of evidence, such as imprecision, heterogeneity (see Section 2.3.11) and publication bias (refer to Section 2.3.13) are also taken into consideration.

Sensitivity analyses

Providing there are sufficient included RCTs or NRSIs, sensitivity analyses (refer to Section 10.14[44] and Section 13.3.5.6 of the Cochrane Handbook[55]) will be performed where meta-analyses have been conducted to explore the influence of the following on treatment effect:

- exclusion of high risk of bias studies (refer to Section 12.4.4.2, Cochrane Handbook 2019[54])
- exclusion of studies with notably long follow-up times or large study populations, to establish how much influence they have on results
- exclusion of cluster RCTs (refer to Section 23.1.6, Cochrane Handbook 2019[31]);
- exclusion of small studies (n<20).

2.3.15 Subgroup analyses

Subgroup analyses (refer to Sections 10.11.2 and 10.11.3 of the Cochrane Handbook 2019[44]) will be undertaken if applicable, if there are sufficient numbers of RCTs or NSRIs and on the basis of:

- relevant sub-populations, such as by disease severity (eg, stage of kidney disease, stage of cancer) and by recognised disease subgroups (e.g. different forms of arthritis, type 1 versus type 2 diabetes). Justification for the use of subgroups is based on the variation of treatment effect as it relates to the biological disease severity or classification.
- intervention-specific groups: multi-modal versus specific single modal interventions.

It is acknowledged that the subgroups will be context-dependent on the populations and modal interventions included in the Review.

2.3.16 Certainty of the evidence

The certainty of evidence will be assessed using the GRADE approach, which involves considering a range of factors that may decrease or increase certainty in the evidence to arrive at an overall 'certainty of the evidence' rating.[36] The certainty will be categorised as:

- **High:** We are very confident that the true effect lies close to that of the estimate of the effect.
- **Moderate:** We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
- **Low:** Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
- **Very low:** We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Certainty of the evidence will be assessed for key outcomes (critical outcomes and important (but not 'not critical') outcomes). RCTs and NRSIs will be assessed separately. Assessment will be by two independent reviewers: disagreements will be resolved by discussion between the two reviewers, with a third reviewer consulted if the former cannot reach agreement.

The GRADE approach commences by first assessing the following five factors and considering whether their certainty should be downgraded:[36]

- **Risk of bias:** as assessed by the risk of bias tools (Section 2.3.6). The overall risk of bias across all studies contributing to each result and the extent to which high risk of bias studies influence the result (i.e. the weight these studies have in the meta-analysis) will be considered.
- **Imprecision:** there is greater imprecision indicated by wide confidence intervals and small sample sizes. Further, imprecision is indicated when the confidence interval crosses the minimal clinically important threshold where the decision between recommending and not recommending a treatment is made and therefore encompasses both benefit and harm. The minimally clinically important difference (MCID) will be sourced from published reports where possible, or will be guided by advice from the NTWC (see section 2.3.8); alternatively, a 25% relative risk increase of 25% or more will used as a default threshold for appreciable harm or benefit (GRADE Handbook section 5.2.4.2). If the MCID is not crossed by the confidence interval, imprecision may also be indicated by optimal information size (OIS). The OIS is calculated as the total number of participants included in the Review for a key outcome that is less than the number required for a sufficiently powered trial.[58] In dichotomous outcomes, imprecision is indicated if the OIS is not met or when the OIS is met and the 95% confidence interval overlaps 'no effect'. An exception to rating down imprecision when the OIS is not met would be where the event rate was very low and the sample size was very large, with at least 2,000 participants (GRADE Handbook section 5.2.4.1). Similar criteria for rating down for imprecision apply to continuous outcomes, including when sample sizes are less than 400 (GRADE Handbook section 5.2.4.2).
- **Inconsistency:** reflected by the heterogeneity of the results. This will involve visual inspection of the overlap in confidence intervals, in combination with cautious interpretation of heterogeneity statistics, and whether any observed inconsistency can be explained.
- **Indirectness:** reflected by how well the studies match all elements of the PICO criteria (see Sections 2.1.1, 2.1.3, 2.1.4 and 2.1.5). This includes how applicable a study's population, intervention and outcomes are to that of the PICO criteria's and whether the outcome results are measured directly (a patient-important outcome) or by a surrogate endpoint (GRADE Handbook section 5.2.3).
- **Publication bias**: as described in Section 2.3.13.

Following the assessment of the first five factors, the certainty of evidence may then be upgraded (i.e. certainty in the evidence may be increased) in the presence of:

- very large effect size
- a dose-response relationship

- increased effect despite having plausible confounders that should have reduced effect, or results show no effect despite plausible confounders that should have led to increased effect.

Incorporating recent advice from the GRADE Working Group regarding the use of ROBINS-I in evaluating risk of bias for NRSIs,[59, 60] the GRADE assessment process will commence at high certainty of evidence for both RCTs and NRSIs.

For each of the GRADE domains (risk of bias, precision, consistency, directness/indirectness and publication and reporting bias), an assessment will be made as to whether there are low or no concerns (no subtractions to the certainty of evidence rank), serious concerns (subtract 1) or very serious concerns (subtract 2). Following an assessment of the factors for downgrading certainty, consideration will be given to whether there are any circumstances that warrant 'rating up' certainty in the evidence, based on the three factors outlined above and in accordance with GRADE guidance.[36] The rationale for downgrading (or upgrading) the certainty in the evidence will be provided as footnotes (see 'Summary of findings' tables).

2.3.17 'Summary of findings' tables

The findings from RCTs and NRSIs will be presented separately within the Summary of Findings table. Findings will be presented as point estimates (mean differences and risk ratios) with 95% confidence intervals, for the key outcomes that were assessed by meta-analysis (see Table 2 for the Summary of Findings table template). For studies that were assessed by vote-counting, the vote-count and non-parametric test results will also be presented in the Summary of Findings table (see section 2.3.12). See Outcome 4 in Table 2 (below) for example of how these would be presented. Where there is sufficient data, a Summary of Findings table will be produced for each population/condition group.

The certainty of the evidence (GRADE) along with any reasons for downgrading will also be presented. The tables will be generated using GRADEPro software.

Table 2: Summary of Findings table template

Summary of findings:

Whole system multi-modal or single modal interventions delivered in the context of naturopathic practice compared to comparator for existing conditions or to prevent a condition a person has risk factors for

Patient or population: people undergoing naturopathic care for existing conditions or to prevent a condition a person has risk factors for
Setting: unrestricted
Intervention: whole system multi-modal or single modal interventions delivered in the context of naturopathic practice
Comparison: comparator

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with comparator	Risk with whole system multi-modal or single modal interventions delivered in the context of naturopathic practice				
Outcome 1	Mean (95% CI, SD or SE)	**Mean (95% CI, SD or SE)**	Mean difference (95% CI)	N participants (N RCTs)	⊕⊕⊕◯[a] Moderate	
Outcome 2	Absolute risk (95% CI, SD or SE)	Absolute risk (95% CI, SD or SE)	Relative risk/odds ratio (95% CI)	N participants (N observational studies)	⊕◯◯◯[a,b,c] Very low	
Outcome 3	[Narrative description of studies with large population samples and low risk of bias that cannot be included in meta-analyses.			N participants (1 RCT/observational study)	⊕⊕◯◯[a] Low	
Outcome 4	[Narrative description for vote-counting analyses – studies that could not be analysed quantitatively]			N studies in favour of intervention N studies favour of comparator	⊕◯◯◯[a,b,c] Very low	Include direction of effect and Sign test

Summary of findings:

Whole system multi-modal or single modal interventions delivered in the context of naturopathic practice compared to comparator for existing conditions or to prevent a condition a person has risk factors for

Patient or population: people undergoing naturopathic care for existing conditions or to prevent a condition a person has risk factors for
Setting: unrestricted
Intervention: whole system multi-modal or single modal interventions delivered in the context of naturopathic practice
Comparison: comparator

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with comparator	Risk with whole system multi-modal or single modal interventions delivered in the context of naturopathic practice				

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

CI: Confidence interval

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

Explanations

a. Reason 1
b. Reason 2
c. Reason 3

Evidence statements will be developed using the certainty of evidence assessed, according to the standardised wording provided by GRADE guideline #26 (Santesso 2020)[61].

Appendix A : References

1. Practitioner Research and Collaboration Initiative (PRACI). PRACI for Researchers: PRACI; 2018. Available from: https://praci.com.au/.
2. Myers SP, Vigar V. The State of the Evidence for Whole-System, Multi-Modality Naturopathic Medicine: A Systematic Scoping Review. *J Altern Complement Med*. 2019;25(2):141-68.
3. World Naturopathic Federation. About Naturopathy: WNF; 2020. Available from: http://worldnaturopathicfederation.org/about-naturopathy/.
4. Natural Therapy Pages. What to Expect in a Naturopathic Consultation 2020 updated 8 October 2020. Available from: https://www.naturaltherapypages.com.au/article/what-to-expect-in-a-naturopathic-consultation.
5. Wellness Centre. Wellness Centre Naturopathic Consultrations 2019. Available from: https://wellnesscentre.net.au/wollongong-naturopath/consultation-process.cfm#:~:text=Naturopathic%20consultations%20are%20scheduled%20and,the%20time%20spent%20in%20consultation.
6. Better Health Channel. Naturopathy: Victoria State Government. Available from: https://www.betterhealth.vic.gov.au/health/ConditionsAndTreatments/naturopathy.
7. Australian Naturopathic Practitioners Association. What is naturopathy? : ANPA. Available from: https://anpa.asn.au/what-is-naturopathy/.
8. PPC Herbs. Visiting a Naturopath: Questions You Need to Ask 2018 updated 19 January 2018. Available from: https://ppcherbs.com.au/visiting-naturopath-questions-need-ask/.
9. World Naturopathic Federation. WNF White Paper: Naturopathic Philosophies, Principles and Theories. Canada: World Naturopathic Federation; 2017.
10. Australian Register of Naturopaths and Herbalists. AROHAH Documents: ARONAH. Available from: http://www.aronah.org/aronah-documents/.
11. Australian Register of Naturopaths and Herbalists. Frequently asked questions: practitioners and students: ARONAH. Available from: http://www.aronah.org/frequently-asked-questions-practitioners/.
12. Ooi SL, McLean L, Pak SC. Naturopathy in Australia: Where are we now? Where are we heading? *Complementary Therapies in Clinical Practice*. 2018;33:27-35.
13. Wardle JL, Sibbritt DW, Adams J. The interface with naturopathy in rural primary health care: a survey of referral practices of general practitioners in rural and regional New South Wales, Australia. *BMC Complementary and Alternative Medicine*. 2014;14(1):238.
14. Ooi SL, Rae J, Pak SC. Implementation of evidence-based practice: A naturopath perspective. *Complementary Therapies in Clinical Practice*. 2016;22:24-8.
15. Oberg EB, Bradley R, Cooley K, Fritz H, Goldenberg J. Estimated effects of whole-system naturopathic medicine in select chronic disease conditions: A systematic review. *Altern Integr Med*. 2015;4(2):1-12.
16. NHMRC. Australian Dietary Guidelines: National Medical Health and Medical Research Council; 2013. Available from: https://www.eatforhealth.gov.au/guidelines.
17. AIHW. Poor diet. Canberra: Australian Institute of Health Welfare; 2019.
18. AIHW. Australian Burden of Disease Study: impact and causes of illness and death in Australia 2015. Canberra: Australian Institute of Health Welfare; 2019.
19. La Torre G, Saulle R, Di Murro F, Siliquini R, Firenze A, Maurici M, et al. Mediterranean diet adherence and synergy with acute myocardial infarction and its determinants: A multicenter case-control study in Italy. *PLOS ONE*. 2018;13(3):e0193360.
20. Boon H, Stewart M, Kennard MA, Guimond J. Visiting family physicians and naturopathic practitioners. Comparing patient-practitioner interactions. *Canadian family physician*. 2003;49(11):1481-7.
21. Oberg EB, Bradley R, Hsu C, Sherman KJ, Catz S, Calabrese C, et al. Patient-reported experiences with first-time naturopathic care for type 2 diabetes. *PLoS One*. 2012;7(11):e48549.
22. Zolnierek KB, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care*. 2009;47(8):826-34.

23. Williams-Piehota PA, Sirois FM, Bann CM, Isenberg KB, Walsh EG. Agents of change: how do complementary and alternative medicine providers play a role in health behavior change. *Altern Ther Health Med*. 2011;17(1):22-30.
24. Lin V, Bensoussan A, Myers SP, McCabe P, Cohen M, Hill S, et al. The Practice and Regulatory Requirements of Naturopathy and Western Herbal Medicine. Victoria: School of Public Health, La Trobe University; 2006.
25. Peltzer K, Pengpid S. Prevalence and determinants of traditional, complementary and alternative medicine provider use among adults from 32 countries. *Chinese Journal of Integrative Medicine*. 2018;24(8):584-90.
26. Jagtenberg T, Evans S, Grant A, Howden I, Lewis M, Singer J. Evidence-based medicine and naturopathy. *Journal of Alternative & Complementary Medicine*. 2006;12(3):323-8.
27. NHAA. Regulation of Herbalists and Naturopaths in Australia. Available from: https://www.nhaa.org.au/mediareleases/be-informed/regulation-of-practitioners.
28. Australasian Cochrane Centre (Monash University). Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies: . Part A. Overview report for naturopathy as a health service. Canberra: National Health and Medical Research Council; 2015.
29. Baggoley C. Review of the Australian Government Rebate on Natural Therapies for Private Health Insurance. Canberra; 2015.
30. Ryan R, Hill S, Broclain D, Horey D, Oliver S, Prictor M. Cochrane Consumers and Communication Review Group. Study Design Guide: Cochrane; 2013 updated June 2013. Available from: http://cccrg.cochrane.org/author-resources.
31. Higgins JPT, Eldridge S, Li T. Chapter 23: Including variants on randomized trials. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
32. Noyes J, Booth A, Cargo M, Flemming K, Harden A, Harris J, et al. Chapter 21: Qualitative evidence. 2019 [cited 14 September 2020]. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane, Accessed 14 September 2020. Available from: www.training.cochrane.org/handbook.
33. Evans SR. Clinical trial structures. *Journal of experimental stroke & translational medicine*. 2010;3(1):8.
34. Reeves BC, Deeks JJ, Higgins JPT, Shea B, Tugwell P, Wells GA. Chapter 24: Including non-randomized studies on intervention effects. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
35. COMET Initiative. Core Outcome Measures in Effectiveness Trials (COMET). Available from: http://www.comet-initiative.org/.
36. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach 2013. Available from: https://gdt.gradepro.org/app/handbook/handbook.html#h.svwngs6pm0f2.
37. Cooley K, Oberg E, Goldenberg JS, Bradley R, Seeley D, Calabrese C, et al. Systematic review of clinical studies of whole practice naturopathic medicine. PROSPERO 2012 CRD42012002176
2012. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42012002176.
38. Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. Chapter 4: Searching for and selecting studies Cochrane, 2019. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
39. Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. Technical Supplement to Chapter 4: Searching for and selecting studies. 2021. In: Cochrane Handbook for Systematic Reviews of Interventions Version 62 (updated Feburary 2021) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
40. Higgins JPT, Lasserson T, Chandler J, D. T, Thomas J, Flemyng E, et al. Standards for the conduct of new Cochrane Intervention Reviews. 2020 March 2020 [cited 30 July 2020]. In: Methodological Expectations of Cochrane Intervention Reviews [Internet]. London: Cochrane, Accessed 30 July 2020. Available from:

https://community.cochrane.org/sites/default/files/uploads/Version%20March%202020%20Final%20Online%20version.pdf.
41. Li T, Higgins JPT, Deeks JJ. Chapter 5: Collecting data. 2019 [cited 4 August 2020]. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane, Accessed 4 August 2020. Available from: www.training.cochrane.org/handbook.
42. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj*. 2014;348:g1687.
43. Higgins JPT, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effect. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
44. Deeks JJ, Higgins JPT, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
45. Sterne JAC, Savović 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.
46. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
47. Risk of Bias 2 Development Group. Revised Cochrane risk-of-bias tool for randomised trials (RoB 2) - Template for completion. 2019.
48. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
49. Sleith C. SIGN Methodology Checklist 4: Case-control studies (version 2.0): Scottish Intercollegiate Guidelines Network; 2012 updated 28 May 2012. Available from: https://www.sign.ac.uk/what-we-do/methodology/checklists/.
50. Cochrane Collaboration. The Risk of Bias In Non-randomised Studies - of Interventions (ROBINS) assessment tool (version for cohort-type studies). 2016.
51. Sterne JAC, Hernán MA, McAleenan A, Reeves BC, Higgins JPT. Chapter 25: Assessing risk of bias in a non-randomized study. 2019 [cited 9 September 2020]. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane, Accessed 9 September 2020. Available from: www.training.cochrane.org/handbook.
52. Review Manager (RevMan) [computer program] Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2014.
53. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019): Cochrane; 2019. Available from: www.training.cochrane.org/handbook.
54. McKenzie JE, Brennan SE. Chapter 12: Synthesizing and presenting findings using other methods. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
55. Page MJ, Higgins JPT, Sterne JAC. Chapter 13: Assessing risk of bias due to missing results in a synthesis. 2019. In: Cochrane Handbook for Systematic Reviews of Interventions version 6 (updated July 2019) [Internet]. Cochrane. Available from: www.training.cochrane.org/handbook.
56. ROB-ME Development Group. Risk of Bias due to Missing Evidence (ROB-ME): a new tool for assessing risk of non-reporting biases in evidence syntheses Melbourne: Cochrane Collaboration; 2020 updated 24 October 2020. Available from: https://sites.google.com/site/riskofbiastool/welcome/rob-me-tool.
57. Boutron I, Page MJ, Higgins JPT, Altman DG, Lundh A, Hróbjartsson A. Chapter 7: Considering bias and conflicts of interest among the included studies. 2019 [cited 30 July 2020]. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane, Accessed 30 July 2020. Available from: www.training.cochrane.org/handbook.

58. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision. *Journal of clinical epidemiology*. 2011;64(12):1283-93.
59. Schünemann 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. *Journal of Clinical Epidemiology*. 2019;111:105-14.
60. Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. 2019 [cited 9 September 2020]. In: Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019) [Internet]. Cochrane, Accessed 9 September 2020. Available from: www.training.cochrane.org/handbook.
61. 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*. 2020;119:126-35.
62. 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. *Res Synth Methods*. 2020.
63. Glanville J, Lefebvre C, Wright K. ISSG Search Filter Resource [Internet] York (UK): The InterTASC Information Specialists' Sub-Group; 2008 updated 4 February 2020. Available from: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home.
64. Furlan AD, Irvin E, Bombardier C. Limited search strategies were effective in finding relevant nonrandomized studies. *J Clin Epidemiol*. 2006;59(12):1303-11.
65. Page MJ, Cumpston M, Chandler J, Lasserson T. Chapter III: Reporting the review. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated August 2019): Cochrane; 2019.

Appendix B : Proposed search strategies

B.1 MEDLINE VIA OVID

In the proposed search strategy for Medline (Table 3), sensitive search filters to identify RCTs from Section 3.6.1 of the Cochrane Handbook's Technical Supplement 4.S1 (2019) were applied.[39] Search strategies to identify NRSIs with control groups were based on Waffenschmidt (2020)[62], cited by the Information Specialists' Sub-Group (ISSG) Search Filter Resource.[63]

Table 3: Proposed search strategy for Medline

#	Search terms
#1	Randomized controlled trial.pt.
#2	Controlled clinical trial.pt.
#3	Randomized.ab.
#4	Placebo.ab.
#5	Drug therapy.fs.
#6	Randomly.ab.
#7	Trial.ab.
#8	Groups.ab.
#9	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8
#10	exp cohort studies
#11	exp Epidemiologic Studies
#12	exp Clinical Trial
#13	exp Evaluation Studies as Topic
#14	exp Statistics as Topic
#15	(control and (group* or study)).mp.
#16	(time and factors).mp.
#17	(program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*).mp.
#18	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17
#19	animals/ not humans
#20	(Editorial or Comment or Letter or Newspaper article).pt.
#21	hi.fs. or case report.mp.
#22	#19 OR #20 OR #21
#23	exp Naturopathy
#24	Naturopath*.tw.
#25	Natural medicine.tw.
#26	Natural therap*.tw.
#27	Naturoceutical*.tw.
#28	Naturopathic.tw.
#29	Integrative Medicine
#30	#23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29
#31	#9 AND #30
#32	#18 AND #30
#33	**#31 NOT #22**
#34	**#32 NOT #22**

Abbreviations: ab, abstract; exp, explode MeSH term; fs, floating subject heading; MeSH, medical subject heading; mp, maps to keyword (mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms); pt, publication type; sh, MeSH heading subject; tw, text word;

Date of search: 7 April 2021
Note: integrative medicine included as proxy for complementary medicine

B.2 EMBASE VIA OVID

In the proposed search strategy for Embase (Table 4), sensitive search filters to identify RCTs from Section 3.6.2 of the Cochrane Handbook's Technical Supplement 4.S1 (2019) were applied.[39] Search strategies to identify NRSIs with control groups were based on both Fixed Methods A and B from Furlan[64] as cited by the ISSG Search Filter Resource.[63] MeSH search terms for naturopathy and therapies were used to identify Emtree synonyms for inclusion in the search strategy.

Table 4: Proposed search strategy for Embase

#	Search terms
#1	Randomized controlled trial
#2	Controlled clinical study
#3	Random$.ti,ab.
#4	Randomization
#5	Intermethod comparison
#6	Placebo.ti,ab.
#7	(compare or compared or comparison).ti.
#8	((evaluated or evaluate or evaluating or assessed or assess) and (compare or compared or comparing or comparison)).ab.
#9	(open adj label).ti,ab.
#10	((double or single or doubly or singly) adj (blind or blinded or blindly)).ti,ab.
#11	double blind procedure
#12	parallel group$1.ti,ab.
#13	(crossover or cross over).ti,ab.
#14	((assign$ or match or matched or allocation) adj5 (alternate or group$1 or intervention$1 or patient$1 or subject$1 or participant$1)).ti,ab.
#15	(assigned or allocated).ti,ab.
#16	(controlled adj7 (study or design or trial)).ti,ab.
#17	(volunteer or volunteers).ti,ab.
#18	human experiment
#19	trial.ti.
#20	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19
#21	(random$ adj sampl$ adj7 (cross section$ or questionnaire$1 or survey$ or database$1)).ti,ab.
#22	comparative study/ or controlled study/ or randomi?ed controlled.ti,ab. or randomly assigned.ti,ab.
#23	#21 NOT #22
#24	Cross-sectional study/ not (randomized controlled trial/ or controlled clinical study/ or controlled study/ or randomi?ed controlled.ti,ab. or control group$1.ti,ab.)
#25	(((case adj control$) and random$) not randomi?ed controlled).ti,ab.
#26	(Systematic review not (trial or study)).ti.
#27	(nonrandom$ not random$).ti,ab.
#28	"Random field$".ab,ti.
#29	(random cluster adj3 sampl$).ti,ab.
#30	(review.ab. and review.pt.) not trial.ti.
#31	we searched.ab. and (review.ti. or review.pt.)
#32	update review.ab.
#33	(databases adj4 searched).ab.

#	Search terms
#34	(rat or rats or mouse or mice or swine or porcine or murine or sheep or lambs or pigs or piglets or rabbit or rabbits or cat or cats or dog or dogs or cattle or bovine or monkey or monkeys or trout or marmoset$1).ti. and animal experiment
#35	Animal experiment/ not (human experiment/ or human/)
#36	#23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35
#37	#20 NOT #36
#38	Case control study/ or cohort analysis/ or controlled study/ or comparative study/ or intermethod comparison/ or major clinical study/ or outcomes research/ or population research/ or prospective study/ or retrospective study/ or treatment outcome/
#39	Clinical article/ or controlled study/ or major clinical study/ or prospective study/ or cohort.mp. or compared.mp. or groups.mp. or multivariate.mp.
#40	#38 OR #39
#41	(Editorial or Comment or Letter or Newspaper article).pt.
#42	Case report.mp.
#43	#35 OR #41 OR #42
#44	#41 OR #42
#45	Naturopathy
#46	Naturopath*.tw.
#47	Natural medicine.tw.
#48	Natural therap*.tw.
#49	Naturoceutical*.tw.
#50	Naturopathic.tw
#51	Integrative medicine
#52	#45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51
#53	#37 AND #52
#54	#40 AND #52
#55	**#53 NOT #44**
#56	**#54 NOT #43**

Abbreviations: ab, abstract; af, all fields; exp, explode; pt, publication type; sh, subject heading; ti, title; tw, text word
Date of search: 7 April 2021

B.3 COCHRANE Central Register of Controlled Trials (via Cochrane library)

Specific search terms for RCTs and NRSIs were not implemented when searching the Cochrane Database, as it is not appropriate for a pre-filtered database (Cochrane Handbook Box C34).[38] As Cochrane uses MeSH terms, these were implemented according to the proposed Medline search strategy, using the Cochrane-specific suffixes for the field codes.

Table 5: Proposed search strategy for CENTRAL Register of Controlled Trials (via Cochrane library)

#	Search terms
#1	MeSH descriptor: [Naturopathy]
#2	(naturopathy):ti,ab,kw
#3	naturopathic
#4	(natural medicine):ti,ab,kw
#5	(natural NEXT therap*):ti,ab,kw
#6	(naturoceutical):ti,ab,kw
#7	**#1 OR #2 OR #3 OR #4 OR #5 OR #6**
#8	**[LIMIT - Cochrane Library Central Register of Controlled Trials]**

Abbreviations: ab, abstract; kw, keyword; MeSH, medical subject heading; ti, title
Date of search: 7 April 2021

B.4 CINAHL VIA EBSCO

In the proposed search strategy for CINAHL (Table 6), sensitive search filters to identify RCTs from Section 3.6.3 of the Cochrane Handbook's Technical Supplement 4.S1 (2019) were applied.[39] A sensitive search filter for NRSIs was not identified for CINAHL and thus CINAHL subheadings and synonyms for the Medline version of the search filter were implemented. MeSH search terms for naturopathy and therapies were mapped to specific CINAHL subject headings and with appropriate field codes, where possible, for inclusion in the search. It should be noted that some MeSH terms, when translated to CINAHL subject headings, could not be exploded. MeSH terms for which there were no CINAHL subject headings were included as keywords.

Table 6: Proposed search strategy for CINAHL

#	Search terms
#1	MH randomized controlled trials
#2	MH double-blind studies
#3	MH single-blind studies
#4	MH random assignment
#5	MH pretest-posttest design
#6	MH cluster sample
#7	TI randomised OR TI randomized
#8	AB (random*)
#9	TI (trial)
#10	MH (sample size) AND AB (assigned OR allocated OR control)
#11	MH (placebos)
#12	PT (randomized controlled trial)
#13	AB (control W5 group)
#14	MH (crossover design) OR MH (comparative studies)
#15	AB (cluster W3 RCT)
#16	MH animals+
#17	MH (animal studies)
#18	TI (animal model*)
#19	S16 OR S17 OR S18
#20	MH (human)
#21	S19 NOT S20
#22	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15
#23	S22 NOT S21
#24	(MH "Prospective Studies+")
#25	(MH "Epidemiological Research+")
#26	(MH "Clinical Trials+")
#27	(MH "Evaluation Research+")
#28	(MH "Statistics+")
#29	TX control AND TX ((group* or study))
#30	TX (program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*)
#31	TX time AND TX factors
#32	S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31
#33	MH animals NOT TX human
#34	PT Editorial OR PT Comment OR PT Letter OR PT Newspaper article
#35	MW hi OR TX case report
#36	S33 OR S34 OR S35
#37	S32 NOT S36
#38	(MH "Naturopathy")
#39	TX naturopath*
#40	TX natural medicine

#	Search terms
#41	TX natural W1 therap*
#42	TX naturoceutical*
#43	(MH "Integrative Medicine")
#44	TI naturopathic
#45	S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44
#46	**S23 AND S45**
#47	**S37 AND S45**

Abbreviations: +, explode search term; " ", indicated key word search (used when MeSH term from Medline did not map to CINHAL subject heading); MH, exact CINAHL subject heading; MW, word in subject heading; PT, publication type; TI, title; TX, all text
Note the CINAHL subject heading for naturopathy, program evaluation, integrative medicine or ' Delivery of Health Care, Integrated' could not be exploded.
Date of search: 8 April 2021

B.5 AMED VIA OVID

For the AMED proposed search strategy, no sensitive search filter for RCTs and NRSIs were identified. The search filters for Medline were adapted using AMED synonyms. MeSH search terms for naturopathy and therapies were mapped to specific AMED synonyms where possible for inclusion in the search. Floating subject heading and text word field codes are not available in AMED and were substituted with subject heading and title/abstract respectively instead. AMED does not have an additional limiter for 'humans'. However, a restriction to human studies is part of the RCT and NRSI search filters.

Table 7: Proposed search strategy for AMED

#	Search terms
#1	randomized controlled trial.pt.
#2	controlled clinical trial.pt.
#3	randomized.ab.
#4	placebo.ab.
#5	drug therapy.sh.
#6	randomly.ab.
#7	trial.ab.
#8	groups.ab.
#9	or/1-8
#10	exp animals/ not humans.sh.
#11	exp cohort studies/
#12	exp Epidemiology/
#13	epidemiologic studies.mp.
#14	exp Clinical trials/
#15	Evaluation Studies.mp.
#16	exp Statistics/ and topic.mp.
#17	(control and (group* or study)).mp.
#18	(time and factors).mp.
#19	(program or survey* or ci or cohort or comparative stud* or evaluation studies or follow-up*).mp.
#20	or/12-20
#21	animals/ not humans/
#22	(Editorial or Comment or Letter or Newspaper article).pt.
#23	case report.mp.
#24	#21 OR #22 OR #23
#25	#10 OR #22 OR #23

#	Search terms
#26	exp Naturopathy/
#27	"naturopath*".ti,ab.
#28	naturopathic.ti,ab.
#29	natural medicine.ti,ab.
#30	"natural therap*".ti,ab.
#31	exp Integrative Medicine/
#32	#26 OR #27 OR #28 OR #29 OR #30 OR #31
#33	#9 AND #32
#34	#20 AND #32
#35	**#33 NOT #25**
#36	**#34 NOT #24**

Abbreviations: ab, abstract; af, all fields; mp, key word; pt, publication type; sh, subject heading; ti, title
Date of search: 7 April 2021

Appendix C : Table templates

Table 8: Characteristics of studies awaiting classification table template[a]

Study ID (year)	Author
Participant description	• Number of participants • Characteristics of participants: including available demographic data such as age and gender • Setting, for example hospital, outpatient, community, research institute, etc • Study eligibility criteria, including diagnostic criteria
Study methods	• Study type/design, for example, parallel, factorial, cross-over, cluster aspects of design for randomised trials, and/or study design features for non-randomised studies. Grey literature type (if NTREAP/NTWC submitted publication) • Single or multicentre study; if multicentre, number of recruiting centres • Duration of study/dates of study • Unit of analysis (e.g. individual participant, clinic, village, body part) • Statistical methods
Intervention	Number of participants (N) Description of intervention, including modalities, dose, method of administration, frequency of administration, who delivered the intervention
Comparator	Number of participants (N) Description of intervention, including modalities, dose, method of administration, frequency of administration, who delivered the comparator
Outcome	• Primary outcomes: description, including measurement method • Secondary outcomes: description, including measurement method
Funding source	Including 'other material support' for study
Conflicts of interest	• Authors' affiliations • Authors' financial relationship • Other potential conflicts of interest, including those declared by the researchers
Comments	Reason this study is awaiting classification

a For studies that were published after the date of the present systematic review search that would otherwise be eligible for inclusion, including studies submitted to NTWC after last date of the systematic review search, and for studies published in languages other than in English that may be eligible for inclusion according to screening at the title-abstract stage (according to the translated citation). A table is to be completed for each study.
Reference Cochrane Handbook (section III3.4.1)[65] and MECIR Manual (section R59)[40]

Table 9: Characteristics of included studies table template[a]

Study ID (year)	Author
Participant description	• Number of participants • Demographic data, including age • Characteristics of participants at the beginning (or baseline) of the study (e.g. age, sex, comorbidity, socio-economic status) • Setting, for example hospital, outpatient, community, research institute, etc • Region(s) and country/countries from which study participants were recruited • Study eligibility criteria, including diagnostic criteria
Study methods	• Study type/design, for example, parallel, factorial, cross-over, cluster aspects of design for randomised trials, and/or study design features for non-randomised studies. Grey literature type (if NTREAP/NTWC submitted publication) • Single or multicentre study; if multicentre, number of recruiting centres • Recruitment and sampling procedures used (including at the level of individual participants and clusters/sites if relevant) • Duration of study/dates of study (see also participant enrolment start/end dates and length of follow-up below) • Details of random sequence generation, allocation sequence concealment, and masking for randomised trials, and methods used to prevent and control for confounding, selection biases, and information biases for non-randomised studies • Unit of analysis (e.g. individual participant, clinic, village, body part) • Statistical methods used if computed effect estimates are extracted from reports, including any covariates included in the statistical model. Include if the method was intention to treat or per protocol. • Method to prevent/address missing data • Likelihood of reporting and other biases
Enrolment start/end dates Length of follow-up	
Intervention[b]	Number of participants (N) 1. Name and description of intervention 2. Description of rationale, theory or goal of the elements essential to the intervention 3. Materials used in the intervention 4. Procedures used in the intervention 5. Intervention provider 6. Modes of delivering the intervention 7. Location where the intervention occurred 8. Timepoints the intervention was delivered, time period, frequency/number of sessions, duration of intervention session, intensity, dosage. 9. Tailoring of the intervention, if the intervention is personalised, titrated or adapted, the rationale and method for doing so; 10. Modifications to intervention, when they occurred, why and how. 11. Strategies to maintain or improve adherence/fidelity to intervention, if assessed. 12. Actual adherence or fidelity to intervention, if assessed
Comparator[b]	Number of participants (N) 1. Name and description of comparator 2. Description of rationale, theory or goal of the elements essential to the comparator 3. Materials used in the comparator 4. Procedures used in the comparator 5. Provider of the comparator 6. Modes of delivering the comparator 7. Location where the comparator was administered 8. Timepoints the comparator was delivered, time period, frequency/number of sessions, duration of comparator session, intensity, dosage. 9. Tailoring of the comparator, if the comparator is personalised, titrated or adapted, the rationale and method for doing so; 10. Modifications to comparator, when they occurred, why and how. 11. Strategies to maintain or improve adherence/fidelity to comparator, if assessed. 12. Actual adherence or fidelity to comparator, if assessed

Study ID (year)	Author
Outcome[c,d]	• Primary outcomes: Description, including measurement method • Secondary outcomes: Description, including measurement method • Whether there is evidence that the outcome domain was assessed (especially important if the outcome was assessed but the results not presented) • Measurement tool or instrument (including definition of clinical outcomes or endpoints); for a scale, name of the scale, upper and lower limits, and whether a high or low score is favourable, definitions of any thresholds if appropriate • Specific metric (e.g. post-intervention anxiety, or change in anxiety from baseline to a post-intervention time point, or post-intervention presence of anxiety (yes/no)) • Method of aggregation (e.g. mean and standard deviation of anxiety scores in each group, or proportion of people with anxiety) • Timing/timepoints of outcome measurements (e.g. assessments at end of eight-week intervention period, events occurring during the eight-week intervention period) • For each group, and for each outcome at each time point: number of participants randomly assigned and included in the analysis; and number of participants who withdrew, were lost to follow-up or were excluded (with reasons for each) • If subgroup analysis is planned, the same information would need to be extracted for each participant subgroup
Funding source	Including 'other material support' for study
Conflicts of interest	• Authors' affiliations • Authors' financial relationship • Other potential conflicts of interest, including those declared by researchers
Reasons for exclusion from data synthesis	For NRSIs identified as at critical risk of bias according to ROBINS-I or were assessed as unacceptable/rejected by the SIGN checklist, state reasons here.

a A table is to be completed for each study.
b Based on the TIDieR checklist[42]
c The Cochrane Handbook also includes 'Adverse outcomes need special attention depending on whether they are collected systematically or non-systematically (e.g. by voluntary report)', but for the present review, safety and adverse events are out of scope.
d Up to 7 pre-specified outcomes identified as critical or as important but not critical to decision-making.
Reference Cochrane Handbook section 5.3.1 (2019)[41]

Table 10: Outcomes data extraction table template for individual eligible studies table template

Study ID Type No. participants	Population / condition	Comparison	Outcome domain	Outcome measure and measurement details	Timepoint	Intervention Results N n/N (%) or mean (SD)	Comparator results N n/N (%) or mean (SD)	Point estimate (RR, OR, MD, SMD) (95% CI), p-value direction of effect	Overall risk of bias[a]
Jones (2020) *Study design: RCT* N=148	*Adults aged 18- 79 (mean age 45) with chronic low back pain*	*Naturopathy (whole-system) vs. placebo*	*Pain*	*VAS (0-100)* *higher score means more pain*	*At end of intervention (12 weeks post-randomisation)*	*N=67* *Mean = 29* *SD = 17*	*N=141* *Mean = 41* *SD = 15*	*MD -12.00 (-20.57, -3.43), p<0.001* *favours naturopathy*	
					At longest follow-up (12 months follow up)	*N=61* *Mean = 6* *SD = 17.18*	*N=134* *Mean = 37* *SD = 17.18*	*MD -31.00 (-42.86, -19.14), p<0.001* *favours naturopathy*	
			Quality of Life	*SF-36* *higher score means better QoL*	*At end of intervention (12 weeks post-randomisation)*				
					At longest follow-up (12 months follow up)				

Abbreviations: MD, mean difference; OR, odds ratio; RR, relative risk; SD, standard deviation; SMD, standardised mean difference;
[a] As assessed by the Cochrane Risk of Bias tool version 2 (RCTs), ROBINS-I tool (NRSIs) or SIGN checklist (case-control studies). See Table 11 to Table 13.

Table 11: Cochrane Risk of Bias 2 tool for RCTs

Study ID	E.g. Jones, M., Boggs, J., Bloggs, P., & Jones, G. (2020) 'Naturopathy for low back pain: a randomised, controlled trial'			
ROB2 Domain	**Signalling questions**	**Outcome**	**Judgement**	**Comment**
DOMAIN 1: Bias arising from the randomisation process	1.1 Was the allocation sequence random? 1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions? 1.3 Did baseline differences between intervention groups suggest a problem with randomisation? (Optional) What is the predicted direction of bias arising from the randomization process?	<Outcome 1>	High risk / Some concerns /Low risk	
		<Outcome 2>	High risk / Some concerns /Low risk	
		<Outcome 3>	High risk / Some concerns /Low risk	
DOMAIN 2: Bias due to deviations from intended interventions	2.1 Were participants aware of their assigned intervention during the trial? 2.2 Were carers and people delivering the interventions aware of participants' assigned intervention during the trial? 2.3 If Y/PY/NI to 2.1 or 2.2, were there deviations from the intended intervention that arose because of the trial context? 2.4 If Y/PY to 2.3, were these deviations likely to have affected the outcome? 2.5 If Y/PY/NI to 2.4, were these deviations from intended interventions balanced between groups? 2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention? 2.7 If N/PN/NI to 2.6, was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomised? (Optional) What is the predicted direction of bias due to deviations from intended interventions?	<Outcome 1>	High risk / Some concerns /Low risk	
		<Outcome 2>	High risk / Some concerns /Low risk	
		<Outcome 3>	High risk / Some concerns /Low risk	
DOMAIN 3: Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised? 3.2 If N/PN/NI for 3.1, is there evidence that the result was not biased by missing outcome data? 3.3 If N/PN to 3.2, could missingness in the outcome depend on its true value? 3.4 If Y/PY/NI to 3.3, is it likely that missingness in the outcome depended on its true value? (Optional) What is the predicted direction of bias due to missing outcome data?	<Outcome 1>	High risk / Some concerns /Low risk	
		<Outcome 2>	High risk / Some concerns /Low risk	
		<Outcome 3>	High risk / Some concerns /Low risk	
DOMAIN 4: Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate? 4.2 Could measurement or ascertainment of the outcome have differed between intervention groups? 4.3 If N/PN/NI to 4.1 and 4.2, were outcome assessors aware of the intervention received by study participants? 4.4 If Y/PY/NI to 4.3, could assessment of the outcome have been influenced by knowledge of intervention received? 4.5 If Y/PY/NI to 4.4, is it likely that assessment of the outcome was influenced by knowledge of intervention received? (Optional) What is the predicted direction of bias in measurement of the outcome?	<Outcome 1>	High risk / Some concerns /Low risk	
		<Outcome 2>	High risk / Some concerns /Low risk	
		<Outcome 3>	High risk / Some concerns /Low risk	

Study ID	E.g. Jones, M., Boggs, J., Bloggs, P., & Jones, G. (2020) 'Naturopathy for low back pain: a randomised, controlled trial'			
DOMAIN 5: Bias in selection of the reported result	5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis? Is the numerical result being assessed likely to have been selected on the basis of the results from: 5.2 multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain? 5.3 multiple eligible analyses of the data? (Optional) What is the predicted direction of bias due to selection of the reported result?	<Outcome 1>	High risk / Some concerns /Low risk	
		<Outcome 2>	High risk / Some concerns /Low risk	
		<Outcome 3>	High risk / Some concerns /Low risk	
OVERALL risk of bias	**(Optional) What is the overall predicted direction of bias for this outcome?**	**<Outcome 1>**	**High risk / Some concerns /Low risk**	
		<Outcome 2>	**High risk / Some concerns /Low risk**	
		<Outcome 3>	**High risk / Some concerns /Low risk**	

Abbreviations: N, no; NI, no information; PN, probably no; PY, probably yes; Y, yes

Table 12: Risk of Bias in Non-randomised Studies of Intervention (ROBINS-I) tool for cohort studies

Study ID				
Domain	**Signalling questions**	**Outcome**	**Judgement**	**Comment**
Domain 1: Bias due to confounding	*Questions relating to baseline confounding only* 1.1 Is there potential for confounding of the effect of intervention in this study? If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding: 1.2. Was the analysis based on splitting participants' follow up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3. 1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8) 1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains? 1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? 1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention? *Questions relating to baseline and time-varying confounding* 1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	<Outcome 1>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 2>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 3>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	

Study ID				
	1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? Optional: What is the predicted direction of bias due to confounding?			
Domain 2: Bias in selection of participants into the study	2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? If N/PN to 2.1: go to 2.4 2.2. If Y/PY to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention? 2.3 If Y/PY to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? 2.4. Do start of follow-up and start of intervention coincide for most participants? 2.5. If Y/PY to 2.2 and 2.3, or N/PN to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases? Optional: What is the predicted direction of bias due to selection of participants into the study?	<Outcome 1>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 2>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 3>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
Domain 3: Bias in classification of interventions	3.1 Were intervention groups clearly defined? 3.2 Was the information used to define intervention groups recorded at the start of the intervention? 3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? Optional: What is the predicted direction of bias due to classification of interventions?	<Outcome 1>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 2>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 3>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
Domain 4: Bias due to deviations from intended interventions	*If studying the effects of assignment to intervention:* 4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice? 4.2. If Y/PY to 4.1: Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome? *If studying the effects of starting and adhering to intervention:* 4.3. Were important co-interventions balanced across intervention groups? 4.4. Was the intervention implemented successfully for most participants? 4.5. Did study participants adhere to the assigned intervention regimen? 4.6. If N/PN to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention? Optional: What is the predicted direction of bias due to deviations from the intended interventions?	<Outcome 1>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 2>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 3>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
Domain 5: Bias due to missing data	5.1 Were outcome data available for all, or nearly all, participants? 5.2 Were participants excluded due to missing data on intervention status?	<Outcome 1>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	

Study ID				
	5.3 Were participants excluded due to missing data on other variables needed for the analysis? 5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions? 5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data? Optional: What is the predicted direction of bias due to missing data?	<Outcome 2>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 3>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
Domain 6: Bias in measurement of outcomes	6.1 Could the outcome measure have been influenced by knowledge of the intervention received? 6.2 Were outcome assessors aware of the intervention received by study participants? 6.3 Were the methods of outcome assessment comparable across intervention groups? 6.4 Were any systematic errors in measurement of the outcome related to intervention received? Optional: What is the predicted direction of bias due to measurement of outcomes?	<Outcome 1>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 2>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 3>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
Domain 7: Bias in selection of reported result	Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome measurements within the outcome domain? 7.2 ... multiple analyses of the intervention-outcome relationship? 7.3 ... different subgroups? Optional: What is the predicted direction of bias due to selection of the reported result?	<Outcome 1>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 2>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
		<Outcome 3>	Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information	
Overall risk of bias judgment	**Optional: What is the overall predicted direction of bias for this outcome?**	**<Outcome 1>**	**Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information**	
		<Outcome 2>	**Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information**	
		<Outcome 3>	**Low risk/ Moderate risk/ Serious risk/ Critical risk/ No information**	

Abbreviations: N, no; PN, probably no; PY, probably yes; Y, yes

Table 13: SIGN 50 methodological checklist for case-control studies

Study ID	
Question	**Judgement**
Section 1: Internal Validity	
In a well conducted case-control study	
1.1 The study addresses an appropriate and clearly focused question.	Yes/ No/ Can't say
Selection of participants	
1.2 The cases and controls are taken from comparable populations.	Yes/ No/ Can't say
1.3 The same exclusion criteria are used for both cases and controls.	Yes/ No/ Can't say
1.4 What percentage of each group (cases and controls) participated in the study?	Cases: Controls:
1.5 Comparison is made between participants and non-participants to establish their similarities or differences.	Yes/ No/ Can't say
1.6 Cases are clearly defined and differentiated from controls.	Yes/ No/ Can't say
1.7 It is clearly established that controls are non-cases.	Yes/ No/ Can't say
Assessment	
1.8 Measures will have been taken to prevent knowledge of primary exposure influencing case ascertainment.	Yes/ No/ Can't say/ Does not apply
1.9 Exposure status is measured in a standard, valid and reliable way.	Yes/ No/ Can't say
Confounding	
1.10 The main potential confounders are identified and taken into account in the design and analysis.	Yes/ No/ Can't say
Statistical analysis	
1.11 Confidence intervals are provided.	Yes/ No
Section 2: Overall assessment of the study	
2.1 How well was the study done to minimise the risk of bias or confounding	High quality (++) Acceptable (+) Unacceptable – reject (0)
2.2 Taking into account clinical considerations, your evaluation of the methodology used, and the statistical power of the study, do you think there is clear evidence of an association between exposure and outcome?	Yes/ No/ Can't say
2.3 Are the results of this study directly applicable to the patient group targeted by this guideline?	Yes/ No
2.4 Notes – Summarise the authors conclusions. Add any comments on your own assessment of the study, and the extent to which it answers your question and mention any areas of uncertainty raised above.	