Project

Western herbal medicines for preventing and treating health conditions: a protocol for an evidence evaluation

Prepared for

National Health and Medical Research Council

NHMRC | Natural Therapies Working Committee
Canberra ACT 2601

CONFIDENTIAL

Research Protocol prepared by
Health Technology Analysts Pty Ltd

March 2021

Contents

Protocol information

Authors

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

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

Contact person

Please send all correspondence to Health Technology Analysts Pty Ltd

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

email: natural.therapies@htanalysts.com.au

Dates

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

History

The NHMRC has been engaged by the Department of Health (Department) to update the evidence underpinning the *2015 Review of the Australian Government Rebate on Natural Therapies for Private Health Insurance* (2015 Review) (1). The seven natural therapies to be reviewed in the first tranche are naturopathy, Pilates, Rolfing, shiatsu, Tai Chi, Western herbalism and yoga. These therapies are among those excluded from the private health insurance rebate as of 1 April 2019.

To support the NHMRC in their evidence review, Health Technology Analysts (HTAnalysts) has been engaged to conduct a systematic review (SR) of the evidence of clinical effectiveness of Western herbal medicines. Eligible reviews submitted from the public, the Natural Therapies Review Expert Advisory Panel (NTREAP) and the Natural Therapies Working Committee (NTWC) will also be included in the evidence evaluation.

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

List of figures

List of abbreviations

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

1 Background

In 2015, a review of Western herbalism as a health service commissioned by NHMRC found no clear evidence demonstrating its efficacy in treating any clinical condition (4, 5). The 2015 Overview was underpinned by an overview of systematic reviews (SRs) that focused solely on the effects of Western herbalism as a health service and were published in the English language between 2008 to May 2013. SRs of the therapeutic effects of individual herbs were excluded, as were SRs of Chinese and Ayurvedic herbal medicines. The 2015 Overview informed the 2015 Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies, which resulted in Western herbalism and 15 other natural therapies being excluded from private health insurance rebates[1].

In this 2020 review, the evidence evaluation will not be limited by publication date and a broader, more comprehensive search of the literature will be undertaken; including individual herbal medicines on List A of the core herbal medicines (see Appendix A) used by the National Herbalists Association of Australia (NHAA), a peak professional association representing appropriately qualified Western herbalists and naturopaths using herbal medicines as their primary treatment modality. Combination herbal medicines that include *at least one* herb from List A, in combination with other herbal ingredients listed on the Therapeutic Goods Administration (TGA) permissible ingredients list will also be included. The updated review will also include studies that assess these core herbal medicines for primary prevention. Like the 2015 Review, SRs evaluating the effectiveness of Chinese and Ayurvedic herbal medicines will be excluded, as these remain outside the scope of the review.

This review will comprise an Overview of Reviews (a SR of SRs), including SRs reporting randomised controlled trials (RCTs) and pseudorandomised controlled trials (pseudo-RCTs). Eligible comparisons will be Western herbal medicines (WHMs) (individual or combination) versus control (further delineated to WHMs versus placebo and WHMs versus no intervention) and WHMs (individual or combination) versus other intervention. See Section 1.2 for the description of WHMs and Section 3.1.3 for information on which WHMs are in scope for the review.

Studies not published in the English language will not be translated, and databases in languages other than English will not be searched.

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

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

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

1.1 Description of condition and setting

Western herbalism is the primary form of herbal medicine utilised in Australia (7). A Western herbalist engages in extemporaneous compounding of herbs for therapeutic purposes for individuals under their care (8). Today, the practice of Western herbalism includes a holistic treatment framework that believes in treating individuals within a wider social, emotional, economical, spiritual and cultural framework and, like naturopathy, adherence to the principle of 'first do no harm' (9). Western herbalists may practice out of various settings including the home, clinical practices and multimodality centres. A survey of Western herbalists in Australia indicated that most practitioners (97.3%) have access to a herbal dispensary within their clinic (10).

Western herbalism is practised for a range of reasons to improve general health and wellbeing, as well as to treat a variety of clinical and preclinical conditions. The current review is not limited to any particular condition or setting (see 3.1.2 Types of participants) and therefore, a concise description of each condition or problem, and the relevant setting, will be provided after conduct of the review.

1.2 Description of intervention

Western herbalism is a traditional system of plant-based medicine derived primarily from Europe, the United Kingdom and North America (9). While medicinal plants from other herbal traditions, such as Traditional Chinese Medicine and Ayurvedic Medicine can be utilised by Western herbalists, the clinical application of Western herbalism is distinct from these traditions.

Western herbal medicine uses plants and plant material to create medicines to help prevent or treat various illnesses. These materials may use some or all parts of a plant such as flowers, roots, stems and rhizomes, fruits and seeds, leaves and bark. WHMs are administered in various preparations including liquid herbal extracts such as tinctures or fluid extracts, oral tablets or capsules, or through topical application, for example, via poultices, creams and pessaries. Most commonly, liquid herbal extracts are prepared using an alcohol solvent, however, glycerol can be used as an alternative, when alcohol-based preparations may not be appropriate (e.g. when prescribing to children). Medicinal herbs can also be extracted in water, and this is commonly referred to as “tea” (9).

In Australia, the regulation of herbal medicines differs depending upon the form, preparation and dosage of the herbal medicine. The TGA regulates some medicinal herbal products (tablets, capsules and liquid extracts), including through a list of permissible ingredients for products listed on the Australian Register of Therapeutic Goods. Others, such as raw plant materials (dried or fresh) used in teas, are unregulated beyond the guidelines applied to all food substances (8, 9); with the exception of herbal medicines listed as a scheduled substance on the Australian poisons standard (11).

A survey of Western herbalists in Australia indicated that the most common preparation of herbs prescribed was liquid extracts (90%), followed by dried preparations such as teas (4.3%) and tablets and capsules (3.8%) (10). These preparations are usually dispensed as either an individualised mixture of one or multiple herbs or dispensed as proprietary formulae such as premanufactured tablets/capsules. Individual consumers also have access to some premanufactured herbal products through pharmacies, supermarkets and health food stores (10).

1.3 How the intervention might work

It is thought that chemical constituents found in plants used for herbal medicine act in a similar manner to pharmaceutical ingredients, noting that some pharmaceutical ingredients were originally derived from plants (e.g. salicylic acid in aspirin). Like pharmaceutical ingredients, it is thought that the chemical constituents in medicinal plants work on a cellular level within the body. However, unlike a pharmaceutical medicine, which often uses purified or manufactured chemical constituents, herbal medicine utilises the 'whole plant' inclusive of the variety of chemical constituents present in its natural form. Western herbalists therefore use unrefined plant extracts (i.e. fluid extracts, teas, creams etc.) containing several different chemical constituents which are thought to work together synergistically, suggesting that the effect of the 'whole plant' is greater than the sum total of the effects of its individual constituents (12). Western herbalists also claim that toxicity is reduced when 'whole plants' are used instead of purified chemical constituents. Western herbalists claim that this synergy also applies to combinations of plants and claim that combining herbs improves clinical efficacy and reduces adverse effects (12).

Western herbalism emphasises the effects of herbs on individual body systems, with aim of treating the underlying cause of disease. Herbs may be used, but not limited to, their supposed anti-inflammatory, haemostatic, expectorant, antispasmodic, immuno-stimulatory etc. properties.

1.4 Why it is important to do this review

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

The 2015 Overview identified no SRs containing evidence evaluating Western herbalism as a health service. The review noted that while there is a large body of research on the effects of individual herbal agents and remedies, the study of the real life practice and outcomes of herbalism as a health service is a relatively new area of research that has yet to be addressed in SRs (4, 5).

The rationale for conducting this Overview of Reviews is to supplement the evidence and guidance used to inform the 2015 Overview of Western herbalism. That is, to identify studies published since, or not included in, the 2015 review and address the evidence gaps noted. This is to ensure recommendations relating to the use of Western herbalism remain relevant and up to date.

2 Objectives

To conduct an Overview of Reviews to evaluate the effectiveness of WHMs in individuals with a described injury, disease, medical condition or preclinical condition. The Overview will compile the evidence from SRs of RCTs and pseudo-RCTs.

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

3 Methods

Methods reported in this protocol are based on that described in the *Cochrane Handbook for Systematic Reviews of Interventions* (13) and relevant sections in the Joanna Briggs Institute Reviewer's manual (14). Covidence (www.covidence.org), a web-based platform for producing SRs, will be used for screening citations and recording decisions made. Covidence is compatible with EndNote and Microsoft Excel, which will be used for managing citations and data extraction, respectively. Where appropriate, RevMan (15) will be used for the main analyses and GRADEpro GDT software (www.gradepro.org) will be used to record decisions and derive an overall assessment of the certainty of evidence for each outcome guided by GRADE methodology (6). The final approved review protocol is to be registered on the international Prospective Register of Systematic Reviews (PROSPERO).

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

3.1 Criteria for considering reviews for this Overview

3.1.1 Types of reviews

Study design

Eligible SRs are those that examine the effectiveness of eligible WHMs (see Section 3.1.3) compared to control (placebo or no intervention) or another intervention.

An Overview of Reviews was selected as the most appropriate methodology, due to the likelihood of sufficient SRs being available to cover a broad range of populations, interventions and outcomes. The aim is to identify and appraise the best available evidence while ensuring the studies included in the review adequately represent the evidence across the full breadth of the PICO (i.e. all eligible comparisons and outcomes for a population or condition).

The primary study of interest is a SR of RCTs (and pseudo-RCTs), with or without a meta-analysis. If the method of randomisation of a primary study included within a SRs is not specifically stated, or not considered strictly random, then the study will be considered to be pseudorandomised. Where a SR includes pseudo-RCTs, these will be considered as eligible along with data from RCTs. Reviews that do not report study eligibility criteria or conduct a comprehensive search of the literature (i.e. searching more than one database) will not be included. These reviews do not meet the minimum criteria to be considered 'systematic' and may not accurately summarise the body of evidence. Eligible reviews that include a single RCT will be included, as will SRs that include both RCTs and nonrandomised studies of an intervention (NRSIs); however, only evidence from the RCTs (and pseudo-RCTs) will be considered. See Section 3.3.11 for information on how meta-analyses from SRs will be handled where they include ineligible studies (e.g. NRSIs) or where they are missing one or more eligible studies.

Additional study designs will not be considered irrespective of whether a SR for an individual herb, herbal combination or condition is identified. This includes individual RCTs or pseudo-RCTs not part of a SR, nonrandomised comparative studies (i.e. nonrandomised experimental trials, cohort studies, case-control studies, interrupted times series), cross-sectional studies and case series with either post-test of pre-test/post-test outcomes. This is because a SR of primary studies (i.e. RCTs or NRSIs) is not feasible given the timeframe and resources. Where a SR is not identified for an eligible population or intervention, this will be noted as an evidence gap.

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

Publication date

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

Reviews published in languages other than English

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

3.1.2 Types of participants

People of any age with any injury, disease, medical condition or preclinical condition are eligible for inclusion. This includes disease prevention in at-risk healthy populations, which is broadly defined as those who are at increased risk of becoming ill or injured based on social, biomedical or behavioural risk factors (16). For the purposes of this review, social risk factors include income, education, employment and social support; biomedical factors include a person's age, genetic make-up, and health status (such as obesity, high blood pressure, high cholesterol, vitamin deficiency); and behavioural factors include a person's lifestyle choices (e.g. alcohol consumption, diet, exercise, tobacco and other drug use, etc.).

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

3.1.3 Types of interventions

Interventions

This Overview of Reviews aims to examine the effectiveness of core individual and combination herbal preparations used by western herbalists in Australia to treat a described injury, or treat or prevent disease, a medical condition or a preclinical condition.

This includes the use of:

- individual herbal medicines on List A (see Appendix A) of the core herbal medicines used by the NHAA for inclusion in the Western herbal medicine curriculum, or
- combination herbal preparations that include at least one herb from List A (see Appendix A) in combination with other herbal medicines listed on the TGA list of permissible ingredients.

Reviews will be included irrespective of whether primary studies have indicated if the intervention is delivered by a certified practitioner.

There are no limits on the type of herbal preparation (i.e. capsule, tablet, liquid extract, tea etc.), however, the herbal preparation must be administered orally, sublingual or be topically applied. SRs will be stratified (where possible) based on the type of herb and how the intervention is prepared (e.g. liquid herbal extracts such as tinctures or fluid extracts, oral tablets or capsules, or topical application, for example, via poultices, creams and pessaries etc.).

Where additional assistance is required regarding eligibility of a combination herbal preparation, a content expert (Dr Erica McIntyre) will be consulted (See Section 3.3.1). SRs that consider a broader question than intended for this Overview of Reviews (e.g. assesses the effectiveness of WHMs

among other interventions) will be included if the SR specifically assesses the effectiveness of WHMs independent of the other included interventions. If only a subset of studies contained within the SR meet the eligibility criteria for this Overview, then only those eligible primary studies as reported in the SR will be considered (see Section 3.3.11).

Restrictions: Individual herbal medicines that are not on List A of the core herbal medicines used by the NHAA. Combination herbal preparations that do not include at least *one* of the NHAA core herbal medicines or that include herbal ingredients that are not listed on the TGA list of permissible ingredients. Combination herbal formulas derived from non-Western herbal medicine traditions (e.g. Chinese, Tibetan, Ayurvedic etc.) and individualised herbal formulas prescribed by therapists from other traditions. Preparations that are administered via injection (i.e. intravenous, intramuscular, subcutaneous). Preparations that contain non-herbal ingredients (i.e. nutraceuticals or pharmaceuticals), doses or administration routes not permitted by the TGA as a 'Listed' complementary medicine. Dietary interventions that are not described in the study as Western herbal medicine are excluded.

Comparators

There are no restrictions on the type of eligible comparators, noting that the analysis will stratify the evidence into three comparisons: (i) placebo; (ii) no intervention, wait list or usual care (unless active); and (iii) other interventions (inclusive of non-WHMs (i.e. Chinese and Ayurvedic formulations) and usual care if considered active). The decision to stratify control into placebo and no intervention comparisons has been made to account for any potential placebo effect which may occur. For instance, while sham interventions are designed to be a placebo, some have demonstrable clinical effects (17).

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

Co-interventions such as diet, education programs, lifestyle modification or medication may be administered simultaneously to the treatment and control group. Reviews that include studies with co-interventions not provided in the context of Western herbalism will be included if all arms of a study receive the same co-interventions (i.e. the effectiveness of the western herbal medicine is not confounded).

Restrictions: Reviews comparing WHMs with other WHMs will be excluded. Where a review includes a mix of herbal medicine comparators (including WHMs and those from other traditions), data will only be extracted for studies with eligible comparators. Where required, clarification will be sought from experts on the NTWC regarding the eligibility of any herbal comparators.

3.1.4 Types of outcome measures

Outcome role

Outcomes will not be used as a criterion for including SRs.

Outcome domains of interest

Outcomes are intended to align with the reasons why patients use the therapy and/or practitioners prescribe the therapy. This includes recovery, rehabilitation, and changes in disease outcomes and

symptoms (e.g. pain, joint range of motion, strength, balance and accepted surrogate outcomes such as HbA1C for diabetes, body mass index for weight gain or loss, lung function tests), health related psychological/behavioural outcomes, health related quality of life, self-reported benefits, symptoms and functional ability, medication use or compliance with conventional medicine treatment; and injury or disease specific prevention outcomes (e.g. falls prevention, smoking cessation).

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

Outcome measures and timepoints of interest

Any effectiveness outcome anticipated to demonstrate a treatment achieves its intended purpose is eligible for inclusion (18, 19). There are no limitations on time points (e.g. short and long term outcomes) or outcome measure (e.g. objective and subjective measures such as clinical and laboratory assessments and patient-reported outcome measures [PROMS], preferably measured using validated tools, are eligible).

As there are a broad range of populations eligible for inclusion in the review, it is not possible to prespecify outcomes. All prespecified outcomes reported in each eligible SR will be listed in the '*Characteristics of included reviews*' tables; however, results will only be extracted for those outcomes identified as critical or important to the Overview. For each identified population, results for a maximum of seven critical (or important) outcomes will be reported in GRADE 'Summary of Findings' tables with corresponding evidence statements (see Section 3.3.17).

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

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

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

3.2 Search methods for identification of reviews

3.2.1 Electronic searches

The literature search strategy (see Appendix B) was developed in Ovid (for Embase, MEDLINE, and Emcare) based on the key element of research question (i.e. the intervention). The search is not limited by population or outcome, but rather by study type; with methodological filters for identifying SRs and exclusions for publication types developed and published previously (20).

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

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

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

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

3.2.2 Other sources

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

Grey literature is excluded, the exception being evidence reviews commissioned by Australian Government bodies and other national or international bodies that are recommended by NTREAP or committee members.

3.3 Data collection and analysis

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

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

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

3.3.1 Inclusion decisions

Reviews identified in the literature searches

Title/abstract screening

Citations (title/abstracts) retrieved by the literature searches will be imported into EndNote and duplicates removed. Citations will then be imported to Covidence (www.covidence.org), an online tool that streamlines the screening and data extraction stages of a SR.

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

Full text screening

Full text articles identified for possible inclusion in the evidence synthesis will be retrieved and assessed for inclusion by one reviewer. A prespecified, hierarchical approach, as outlined in Appendix C, will be used to annotate reasons for exclusion, with the results of the study selection process illustrated in a PRISMA diagram. Ineligible SRs will be marked with a reason for exclusion and listed in a table in the technical report under '*Characteristics of excluded reviews*'. Where there is uncertainty regarding inclusion, a decision will be made through discussion with the lead reviewer. The lead reviewer will also reinspect a random 20% sample of articles marked as excluded to ensure adherence to the *a priori* exclusion criteria and any differences will be resolved by discussion. If additional expertise or advice regarding the application of the PICO criteria is required, further follow up with the NTWC will occur (noting that NTWC will be presented with excerpts from the publication relevant to the query while remaining blinded to other identifying details regarding the review, included studies or results).

To confirm combination herbal preparations are representative of WHM in Australia, a list of potentially relevant SRs will be supplied to a content expert for independent full text screening. The expert will confirm the appropriateness of the herbal combination/s as meeting the WHM eligibility criteria after examination of the SR (or primary studies within). Advice regarding the relevant grouping or subgrouping of the studies for analysis (with regards to the intervention) will also be sought at this time.

If a review does not contain the required PICO information for a decision to be made regarding eligibility, the information will be sought from the SR authors through an open-ended request. Author names and included primary studies will be used to identify multiple reports arising from the same review. Eligible studies that are not available in English will be noted and managed as described below under '*Reviews published in languages other than English*'.

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

Evidence provided through the Department's public call for evidence

Potentially relevant SRs identified by the NTWC, NTREAP, and other key stakeholders will be considered for inclusion if they satisfy the eligibility criteria described in Section 3.1 above.

All of the submitted literature will be collated, tabulated, and cross-referenced with the evidence identified in the literature search described in Section 3.3.1. In-scope SRs not identified in the literature search will be incorporated into the evidence evaluation. A rationale for exclusion (as noted in Appendix C) will be provided for all SRs and primary studies considered out of scope (documented in a table within the technical report). Overviews provided through the Department's public call for evidence will be checked for eligible SRs.

Potentially relevant RCTs provided through the Department's public call for evidence (see Section 3.2.2) will be checked to confirm if they have been identified within eligible SRs. If the RCT has been published after the literature search date of an eligible SR, they will be noted and discussed under the relevant section of the Overview. Relevant RCTs that are not be associated with a condition or intervention identified in this Overview will be listed in the technical report; however, these studies will not be subject to a formal evidence evaluation.

Reviews published in languages other than English

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

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

The potential risk of language bias and its implications for the Overview will be discussed in relevant sections of the Evaluation Report (such as '*Overall completeness and applicability of evidence*' and

'*Agreements and disagreements with other studies or reviews*'). When assessing the extent to which language bias might influence the conclusions, we will also consider whether English-language reviews included in the Overview searched for, and included, studies published in languages other than English.

3.3.2 Data collection process

Data from systematic reviews

The characteristics of all included SRs will be extracted by one reviewer using a standard pre-tested data extraction and coding form (see Appendix E). Outcome data will be extracted after agreement has been reached regarding the critical and important outcomes to be appraised (see Section 3.1.4). Pretesting will involve all reviewers data extracting the required information from the same three SRs which will be specifically selected to cover the breadth of the PICO identified for inclusion in the review. The lead reviewer will compare the data extraction forms to ensure the relevant data are extracted consistently between reviewers and as planned, with any necessary revisions made to ensure consistency.

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

Data from included studies

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

3.3.3 Requests for data

Eligible SR protocols, SRs published as conference abstracts, and SRs not published in English will be identified for inclusion. Authors will be contacted through an open-ended request for data or further information. If no further data are available, the review will be noted as '*Ongoing*' or '*Reviews awaiting classification'* and will not be included in the evidence appraisal.

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

3.3.4 Data items

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

Included primary studies will also be listed by author, date of publication and eligibility for inclusion in this review. When evaluating the effectiveness of WHM across the same PICO (See Section 3.3.11

and 3.3.12), overlaps and omissions between SRs will be noted in a matrix (see Figure 1). Where data from a selected SR is augmented with data obtained from another SR this will be documented using footnotes.

3.3.5 Missing data

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

3.3.6 Tools to assess risk of bias

The methodological quality of included SRs will be assessed using the AMSTAR-2 quality assessment checklist (21). The AMSTAR-2 consists of 16 domain questions (see Appendix D) that are answered as 'yes', 'no', or 'partial yes'; with a 'yes' answer denoting a positive result. Any notable strengths or limitations of the SR (in reference to the relevant AMSTAR-2 domains) will be reported. If the SR is broader in scope than the clinical question posed in this Overview (i.e. includes other interventions or NRSIs no eligible for inclusion), the overall quality of the SR will be assessed.

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

Risk of bias of included studies

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

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

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

3.3.7 Risk of bias assessment process

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

For each review, we will report our judgement for each item on the AMSTAR-2 checklist (i.e. answer 'yes', 'no', or 'partial yes') (see Appendix D). Limitations of each SR (including a rationale for

judgements with supporting information) will be described in the '*Characteristics of included reviews'* table (see Appendix E).

To ensure consistency among reviewers, pretesting of AMSTAR-2 assessments will be achieved by all reviewers completing assessments for three SRs selected to cover the breadth of the PICO. The lead reviewer will inspect the forms to ensure consistency, and any differences will be resolved through discussion.

3.3.8 Measures of effect

For SRs that do not report a meta-analysis for a relevant comparison because it is not appropriate to do so, effect estimates for the primary studies will be reported (where possible) as described below. If appropriate to meta-analyse, analysis will be performed as per Section 223.3.11.

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

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

3.3.9 Unit-of-analysis issues

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

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

3.3.10 Studies with more than two intervention groups

SRs that have included studies with multiple treatment groups, will be noted in the results table along with a footnote describing how the SR dealt with the multiple treatment groups in their evidence synthesis (e.g. combining like groups to create a single pairwise comparison, double counting the placebo group in a meta-analysis). The implications of the multiple treatment groups will be considered when interpreting the evidence, with any important implications for interpreting results documented in footnotes to the summary of findings table.

3.3.11 Meta-analysis

Synthesis will only be undertaken for SRs that compare WHMs with 'control' (stratified into 'placebo' or 'no intervention'). Pooled results data from SRs comparing WHMs with 'other' interventions will be extracted and presented in data tables (see Appendix E) that include the studies contributing data and the overall AMSTAR-2 assessment, but will not be synthesised further, except where requested

by the NTWC. Where pooled data are not available, individuals study results will be extracted (see Section 3.3.8). These data will be presented as an 'evidence inventory' to provide a snapshot of the available evidence comparing WHMs with 'other' interventions.

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

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

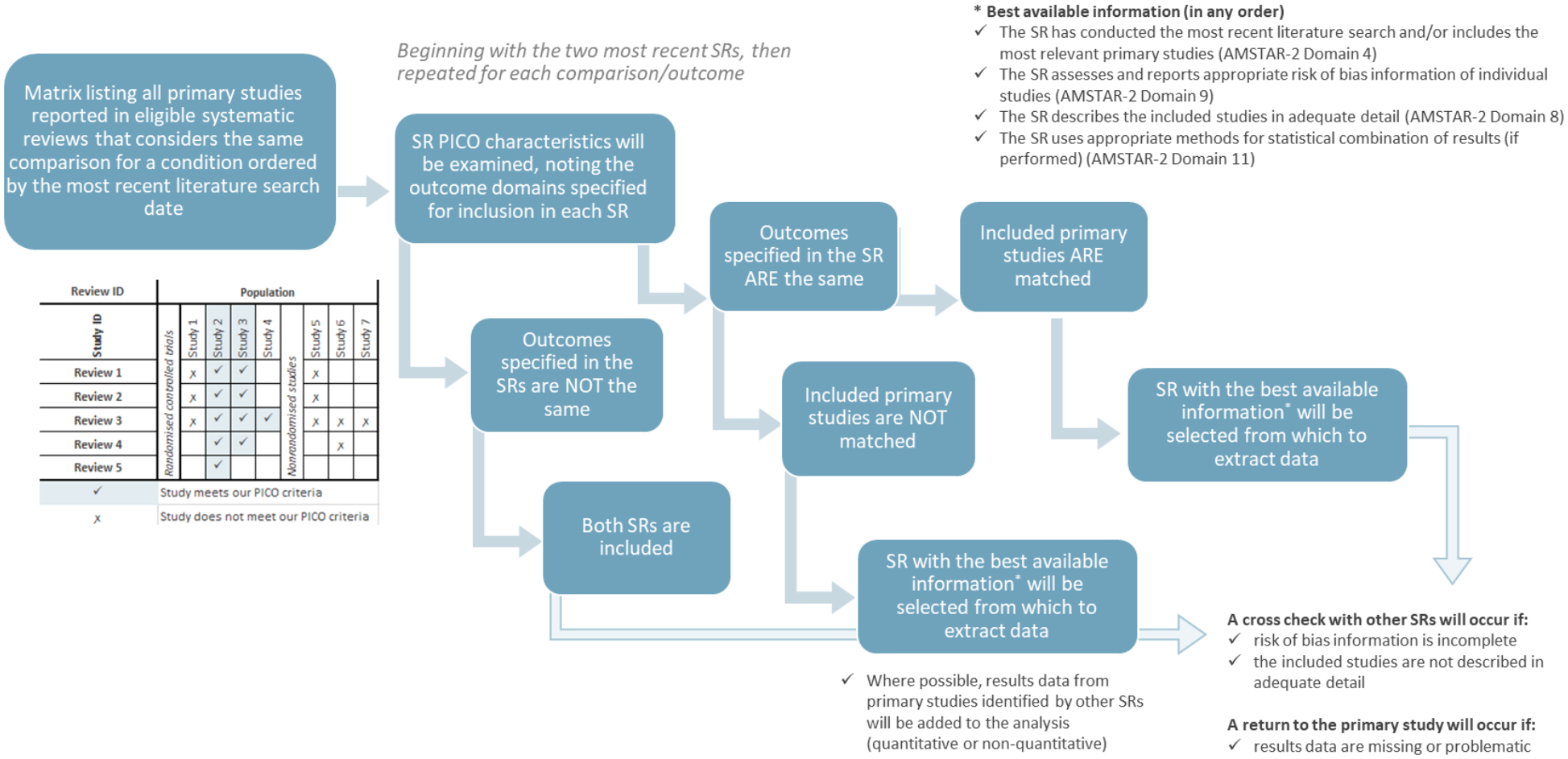
Such cases will be identified by the NTWC through blinded discussions with the evidence reviewer at the data synthesis stage, or prior to provision of the first draft evaluation report.

Data from systematic reviews

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

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

Figure 1 Framework for selecting the SR from which to extract data for any given PICO



Where the selected SR result does not include all eligible studies for a given comparison and outcome, the meta-analysis reported by the selected SR will be updated and re-analysed (where possible). The decision to re-analyse pooled data will be determined by whether:

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

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

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

If the best available SR does not report a meta-analysis for a relevant comparison, and it is appropriate to do so, data synthesis will be performed using RevMan 5.4 and forest plots presented (see Section 3.3.8). Within each comparison we will combine effect estimates across studies for each outcome using a random effects model to take into account expected differences between studies. Statistical heterogeneity will be assessed by visually inspecting the overlap of confidence intervals on the forest plots, formally testing for heterogeneity using the Chi^2 test (using a significance level of α=0.1), and quantify heterogeneity using the I^2 statistic (22).

For SRs where the meta-analysis or primary study results are incompletely reported (e.g. no effect estimate is reported, but the direction of effect is reported along with a *p*-value), we will report the available information (see Section 3.3.8). If the reported information allows for calculation of effect estimates or imputation of missing statistics (e.g. SD), we will perform the calculations as described in Chapter 6 of the Cochrane Handbook (23).

3.3.12 Summary and synthesis when meta-analysis is not possible

The evidence review will provide a structured summary of the results for each condition identified, in tables structured by intervention (further delineated to individual herbs or combinations of herbs), comparator ('placebo', 'control', or 'other' intervention), and outcome domain. Where possible, a visual representation of the results of included studies will be presented in a forest plot (without a summary estimate) grouped by study design features and risk of bias.

The narrative summary will include a brief description of the condition and reviews identified (including review criteria for inclusion or exclusion of studies, population demographics and other key features). Any notable weaknesses within a review, or inconsistency across reviews will be

recorded. This will be followed by a summary of results grouped by intervention, comparator and outcome domain.

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

If there are several eligible SRs identified that evaluate the effectiveness of WHMs across the same PICO, results will be reported from the selected SR based on pre-specified criteria (see Figure 1). In the absence of supplementary quantitative data, results from additional studies identified in other SRs will be described, with the range and distribution of observed effects noted. To describe the overall effect, a simple vote count based on direction of effect will be used (e.g. XX RCTs were identified in the by SR by YY, who reported a pooled effect favouring the intervention for outcome ZZ (MH fixed effects; effect size; 95%CI; p-value, I^2; GRADE). One additional RCT was identified (by SR) that also showed an effect favouring the intervention, however we were unable to add this to the quantitative synthesis because [reasons…]). Any important features of the additional studies (e.g. risk of bias. study design, size) that may influence the interpretation of results will be called out and highlighted in the text.

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

3.3.13 Risk of reporting bias across studies

As noted in Section 3.3.5, the implications for reporting bias within reviews will be considered when interpreting the evidence. Judgements regarding reporting bias across studies will be based on that reported within the SRs (e.g. publication bias, small study effects, other reporting biases). Additional approaches for assessing bias due to missing studies (such as additional searching of clinical trial registers, grey literature, or other reports) will not be performed.

Judgements concerning reporting bias across reviews will be based on available information (e.g. from '*Studies awaiting classification*' etc.) and discussed under '*Overall completeness and applicability of evidence*'. Any variability across reviews and any important flaws in individual reviews will be discussed under *'Quality of the evidence'*.

3.3.14 Addressing risk of bias

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

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

No formal assessment specifically addressing the risk of bias within a SR will be conducted; however, where there are concerns related to the methodological quality of a SR (e.g. due to concerns with study eligibility criteria, methods used to identify, select, or appraise studies, or concerns with interpretation of findings) we will attempt to mitigate the potential bias by cross-checking data

across SRs and re-analysing and re-interpreting results. A brief statement about the impact of any changes made to the evidence reported by a SR (e.g. removal of a study due to inappropriate inclusion, change to the risk of bias assessment for that study, update on the data reported by the SR because they reported an incorrect number) on the overall conclusions of the Overview will be included under the relevant sections of the report (including the '*Overall completeness and applicability of evidence*').

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

3.3.15 Subgroup analyses

We do not plan to undertake any subgroup analyses of subsets of participants within SRs. If, for a particular PICO, there is inconsistency between SR conclusions, we will explore eligibility criteria of the reviews as well as population, intervention and comparator characteristics (e.g. liquid herbal extracts, oral tablets or capsules, or topical application) within the included studies to provide a hypothesis for future reviews.

3.3.16 Certainty of the evidence

Across each population and intervention, we will assess the certainty of the evidence for each critical (or important) outcome using the GRADE approach (6). Only evidence comparing WHMs with 'placebo' and 'no intervention' will be presented (in separate 'Summary of Findings' tables).

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

- *Risk of bias.* Based on the summary assessment of bias across studies (as reported in the priority SR, or supplementary SRs) for each outcome reported for a comparison (24).
- *Inconsistency.* Based on heterogeneity in the observed intervention effects across studies that suggests important differences in the effect of the intervention and whether this can be explained (25).
- *Imprecision.* Based on interpretation of the upper and lower confidence limits in relation to a clinically important threshold (i.e. the confidence interval includes both appreciable benefit and harm); and whether the optimal information size has been reached (i.e. the total number of patients meets the required sample size for a sufficiently powered individual study). In the absence of a published clinically important threshold a rough guide will be used (i.e. a 25% relative risk reduction or increase) (26).
- *Indirectness.* Based on important differences between the review questions and the characteristics of included studies that may lead to important differences in the intervention effects (27).
- *Publication bias.* Based on the extent to which the evidence is available. Publication bias would be suspected when the evidence is limited to a small number of small trials (28).

The certainty of evidence will be categorised as follows:

- High (⊕⊕⊕⊕): further research is very unlikely to change the confidence in the estimate of effect.

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

For each domain, a judgement will be made about whether there are serious, very serious or no concerns; resulting in an overall GRADE describing the certainty of evidence for each outcome. Footnotes will be used to record judgements made about downgrading (or upgrading) the evidence (see Section 3.3.17). Scoring of the certainty of the evidence will begin as 'high' for RCTs (score=4), which can be downgraded by –1 for each domain with serious concerns or –2 for very serious concerns (6).

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

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

Where the selected SR has reported a GRADE certainty of evidence for a PICO, we will replicate that information in the Overview, adjusting for any important differences between the eligibility criteria of the SR and the Overview (i.e. we will assess indirectness of the evidence in relation to the overview PICOs). If the PICO from the selected SR has been updated with supplementary data or has been re-analysed through the removal of studies that do not meet the review PICO (see Section 3.3.11) the GRADE assessment (if available) will be revised to reflect the totality of the evidence. If the selected SR has not evaluated the certainty of the evidence using a GRADE, a *de novo* GRADE assessment will be performed.

In the absence of enough information from the selected SR (or other SRs) to inform our judgements for a certain factor (e.g. risk of bias, imprecision), the certainty of the evidence will be determined by the following system (adapted from the *NHMRC Report on the Evidence: promoting social and emotional development and wellbeing of infants in pregnancy and the first year of life*) (29):

- *Where information regarding one domain is missing and there are no apparent concerns or reasons to downgrade.* GRADE certainty of the evidence will be presented as a range including the rating based on the available information and one level lower to account for uncertainty in the rating (e.g. an outcome has serious concerns due to risk of bias, but there is no information on imprecision, then the overall certainty of evidence will be judged to be moderate to low quality).
- *Where information regarding one domain is missing and there is some indication that the factor with missing information should be downgraded*: The overall quality certainty of

evidence will be downgraded based on the available information, and presented as 'assumed' in brackets. Footnotes will be used to record judgements made about downgrading the evidence in the absence of information.
- *Where information regarding two features is not reported*: the overall quality of the evidence will be judged to be 'unclear'.

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

3.3.17 'Summary of findings' table

For each population, findings for the critical and important outcomes (see Section 3.1.4) will be reported in 'Summary of Findings' tables that will be prepared using the GRADEpro GDT software (www.gradepro.org). The findings comparing WHMs with 'placebo' and 'no intervention' will be presented separately. Estimates of treatment effects for each outcome will be reported as absolute and relative risks (or standardised means). In the absence of quantitative data, a narrative synthesis will be provided (see Section 3.3.12). All critical or important outcomes will be reported, regardless of whether the findings demonstrate a clinically meaningful change.

The 'Summary of Findings' tables will provide a summary of each of the included critical or important outcomes and the certainty of evidence rating for each outcome in a quick and accessible format (6). As part of the 'Summary of Findings' table, an evidence statement pertaining to each outcome will be included. This statement will be guided by the following format:

The use of Western herbal medicines in [population] [is suggested to, may, results] in [little to no effect, reduce, increase, promote etc.] on [outcome] compared with [placebo or no intervention].

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

- the methodology used to identify the evidence base (documented systematic literature search, inclusion and exclusion criteria described),
- the characteristics of the included reviews, including data extraction and AMSTAR-2 assessment forms, and
- detailed results, presented by population and outcome, containing complete information about the evidence assessment.

Contributions of authors

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

NTREAP and NTWC provide expert advice, especially in relation to intervention, study design and eligibility criteria. A methodological review of the draft protocol was conducted by Health Research Consulting (hereco).

Declarations of interest

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

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

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

References

1. Australian Government Department of Health. The 2015 Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies 2019. Available from: https://www.health.gov.au/internet/main/publishing.nsf/Content/phi-natural-therapies.
2. Higgins J, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, et al. Methodological Expectations of Cochrane Intervention Reviews. London: Cochrane; 2019. Available from: https://community.cochrane.org/mecir-manual.
3. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4:1.
4. The Australasian Cochrane Centre. Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies. Part A - Overview report for herbalism as a health service Monash University2014.
5. The Australasian Cochrane Centre. Review of the Australian Government Rebate on Private Health Insurance for Natural Therapies. Part B - Consideration of stakeholder submissions for herbalism as a health service. Monash University 2014.
6. Schünemann H, Brożek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach [Internet]. 2013; (Updated October 2013). Available from: https://gdt.gradepro.org/app/handbook/handbook.html.
7. Hawrelak JA, Wohlmuth H, Pattinson M, Myers SP, Goldenberg JZ, Harnett J, et al. Western herbal medicines in the treatment of irritable bowel syndrome: A systematic review and meta-analysis. Complement Ther Med. 2020;48:102233.
8. Lin V, McCabe P, Bensoussan A, Myers S, Cohen M, Hill S, et al. The practice and regulatory requirements of naturopathy and western herbal medicine in Australia. Risk Manag Healthc Policy. 2009;2:21-33.
9. National Herbalists Association of Australia. Review of the Australian Government rebate on private health insurance for Natural Therapies. 2013 [Available from: https://www.nhaa.org.au/docs/Submissions/NHAA_Private_Health_Insurance_Submission__29_01_2013_.pdf.
10. Casey MG, Adams J, Sibbritt D. An examination of the prescription and dispensing of medicines by Western herbal therapists: a national survey in Australia. Complement Ther Med. 2007;15(1):13-20.
11. Therapeutic Goods Administration. Schedule 2 - Standard for the uniform scheduling of medicines and poisons. No. 27. Regulatory Engagement and Planning Branch, Woden, ACT: Australian Government Department of Health; February 2020. Available from: https://www.legislation.gov.au/Details/F2020L00017/Html/Text#_Toc28084707.
12. Vickers A, Zollman C, Lee R. Herbal medicine. West J Med. 2001;175(2):125-8.
13. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019): Cochrane; 2019. Available from: www.training.cochrane.org/handbook.
14. Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L. Chapter 3: Systematic reviews of effectiveness. In: Aromataris E, Munn Z, editors. Joanna Briggs Institute Reviewer's Manual: The Joanna Briggs Institute; 2017.
15. Review Manager (RevMan) [Computer program]. Version 5.4. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration; 2020.
16. Australian Institute of Health and Welfare. Australia's health 2016. Canberra; 2016.
17. Brim RL, Miller FG. The potential benefit of the placebo effect in sham-controlled trials: implications for risk-benefit assessments and informed consent. J Med Ethics. 2013;39(11):703-7.
18. Merlin T, Weston A, Tooher R. Extending an evidence hierarchy to include topics other than treatment: revising the Australian 'levels of evidence'. BMC Med Res Methodol. 2009;9:34.

19. National Health and Medical Research Council. NHMRC Additional levels of evidence and grades for recommendations for developers of guidelines. 2009. Available from: https://www.mja.com.au/sites/default/files/NHMRC.levels.of.evidence.2008-09.pdf.
20. National Blood Authority. Patient Blood Management Guidelines: Module 6 Neonatal and Paediatrics. Technical report - Volume 2. Canberra, Australia: National Blood Authority; 2016. Available from: https://www.blood.gov.au/pbm-module-6.
21. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.
22. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.
23. Higgins JPT, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 6,0 (updated July 2019).
24. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol. 2011;64(4):407-15.
25. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. J Clin Epidemiol. 2011;64(12):1294-302.
26. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol. 2011;64(12):1283-93.
27. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol. 2011;64(12):1303-10.
28. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. J Clin Epidemiol. 2011;64(12):1277-82.
29. Australian Research Centre for the Health of Women and Babies (ARCH). Evaluation of evidence on the effectiveness of interventions for caregiving practices and behaviours for optimal social and emotional development of infants: an overview of systematic reviews. https://www.nhmrc.gov.au/file/5391/download?token=k6mJopkU The University of Adelaide; 2017.
30. NHAA. Naturopathy and Western Herbal Medicine Course Accreditation Standards Curriculum Mapping Document v1.0. : Naturopaths & Herbalists Association of Australia; Last Modified: 29 May 2018. Available from: https://www.nhaa.org.au/docs/2018_CAS_Mapping_-_FINAL.pdf.

Appendix A – List of core herbal medicines

BINOMINAL NAME	COMMON NAME (Part specified)
Achillea millefolium	Yarrow
Actaea racemosa	Black cohosh (root)
Aesculus hippocastanum	Horse chestnut (seed)
Albizia lebbeck	Albizia
Allium cepa	Onion
Allium sativum	Garlic
Aloe spp.	Aloe
Althaea officinalis	Marshmallow
Andrographis paniculata	Andrographis
Angelica archangelica	Angelica (root)
Apium graveolens	Celery
Arctium lappa	Burdock (root)
Arctostaphylos uva-ursi	Bearberry
Armoracia rusticana	Horseradish
Artemisia absinthium	Wormwood
Astragalus membranaceous	Astragalus
Avena sativa	Oats
Bacopa monniera	Bacopa
Berberis vulgaris	Barberry
Boswellia serrata	Boswellia
Bupleurum falcatum	Bupleurum
Calendula officinalis	Calendula
Camellia sinensis	Tea, Green
Capsicum minimum	Cayenne
Cassia angustifolia	Senna, Indian
Centella asiatica	Gotu kola
Chelidonium majus	Celandine (herb)
Cinnamomum zeylanicum / C.cassia	Cinnamon (bark)
Coleus forskohlii	Coleus
Commiphora myrrha	Myrrh
Crataegus oxyacantha / C.monogyna	Hawthorn
Crocus sativus	Saffron
Curcuma longa	Turmeric (root)
Cynara scolymus	Artichoke, Globe (leaf)
Dioscorea villosa	Wild yam
Drosera rotundifolia / D. angelica / D. intermedia	Sundew
Echinacea spp.	Echinacea

BINOMINAL NAME	COMMON NAME (Part specified)
Eleutherococcus senticosus	Siberian Ginseng
Elytrygia repens	Couch grass
Equisetum arvense	Horsetail (herb)
Eschscholzia californica	California poppy
Eucalyptus globus/ E. spp.	Eucalyptus
Eupatorium perfoliatum	Boneset
Euphorbia hirta	Asthma weed
Euphrasia officinalis	Eyebright
Filipendula ulmaria	Meadowsweet
Frangula purshiana	Cascara
Fucus vesiculosus	Bladderwrack
Galega officinalis	Goat's rue (herb)
Galium aparine	Cleavers
Gentiana lutea	Gentian (root)
Geranium maculatum	Cranesbill
Ginkgo biloba	Ginkgo (leaf)
Glycyrrhiza glabra	Licorice (root)
Gymnema sylvestre	Gymnema
Hamamelis virginiana	Witch Hazel (leaf and bark)
Harpagophytum procumbens	Devil's claw
Hedera helix	Ivy (English - leaf)
Hemidesmus indicus	Hemidesmus
Humulus lupulus	Hops
Hydrastis canadensis	Goldenseal
Hypericum perforatum	St John's wort
Inula helenium	Elecampane (root)
Iris versicolor	Blue flag
Lavandula officinalis / L. angustifolia	Lavender (flower)
Leonurus cardiaca	Motherwort
Linum usitatissimum	Linseed (aka Flaxseed)
Lycopus virginicus	Bugleweed / Gypsyweed
Marrubium vulgare	White Horehound (herb)
Matricaria recruitica	Chamomile (flower)
Melaleuca alternifolia	Tea Tree (oil)
Melissa officinalis	Lemon balm
Mentha x piperita	Peppermint
Nigella sativa	Black cumin
Olea europaea	Olive (leaf)
Paeonia officinalis	Peony (flower and root)

BINOMINAL NAME	COMMON NAME (Part specified)
Panax ginseng	Ginseng (root)
Passiflora incarnata	Passionflower (herb)
Phytolacca decandra / P.americana	Poke root
Pimpinella anisum	Aniseed /Anise (seed)
Piper methysticum	Kava kava
Piscidia erythrina	Jamaican dogwood
Plantago lanceolata	Ribwort
Plantago ovata	Psyllium
Polygonum aviculare	Knotweed (herb)
Prunus serotina	Wild cherry (bark)
Ptychopetalum olacoides	Muira puama / Potency wood
Rehmannia glutinosa	Rehmannia
Rhodiola rosea	Rhodiola
Rosmarinus officinalis	Rosemary (leaf)
Rubus idaeus	Raspberry (leaf)
Rumex crispus	Yellow dock
Salix alba	White Willow (bark)
Salvia officinalis	Sage (leaf)
Sambucus nigra	Elder (flower)
Schisandra chinensis	Schisandra
Scutellaria baicalensis	Baikal Skullcap
Scutellaria lateriflora	Skullcap
Serenoa serrulata / S. repens	Saw Palmetto (berry)
Silybum marianum	St Mary's Thistle
Solidago virgaurea	Goldenrod
Stellaria media	Chickweed
Tanacetum parthenium	Feverfew
Taraxacum officinale	Dandelion
Thuja occidentalis	Thuja
Thymus vulgaris	Thyme
Tilia spp.	Lime (flower)
Tribulus terrestris	Tribulus
Trifolium pratense	Red clover
Trigonella foenum-graecum	Fenugreek (seed)
Turnera diffusa	Damiana (leaf and herb)
Ulmus rubra	Slippery elm
Urtica dioica	Nettle
Vaccinium macrocarpon	Cranberry
Vaccinium myrtillus	Bilberry (fruit)

BINOMINAL NAME	COMMON NAME (Part specified)
Valeriana officinalis	Valerian (root)
Verbascum thapsus	Mullein (flower)
Verbena officinalis	Vervain (root)
Viburnum opulus	Cramp bark
Vitex agnus-castus	Chaste tree (fruit)
Withania somnifera	Withania
Zanthoxylum clava-herculus / Z. americanum	Prickly ash
Zea mays	Corn (silk)
Zingiber officinale	Ginger (root)
Zizyphus jujuba / Z. spinosa	Chinese date

Source: Naturopaths & Herbalists Association of Australia (NHAA) (30) Available at: https://www.nhaa.org.au/docs/2018_CAS_Mapping_-_FINAL.pdf

Appendix B – Literature search strategy

Concept: Study design limits (systematic reviews, not animals, not Chinese or ayurvedic)

1. exp meta analysis/ or meta analysis.mp. or exp systematic review/ or systematic review.mp. or pooled analysis.mp. or ((exp review/ or review.mp.) and (systemat* or pool*).mp.)
2. case report/
3. (editorial or letter or comment or historical article).pt.
4. (animals/ or nonhuman/) not humans/
5. 2 or 3 or 4

Concept: MeSH terms

6. *herbal drugs/
7. *herbaceous agent/
8. *herbal drug/
9. *herbal medicinal product/
10. *medicinal plant/
11. *traditional medicine/
12. *plant extracts/
13. *plants medicinal/
14. *herbalism/
15. *herbal medicine/
16. *phytotherapy/
17. or/6-16
18. 1 and 17
19. (chinese or ayurved$).ti.
20. 18 not (5 or 19)

Concept: individual herbs

21. (((a or achillea) adj millefoli*) or yarrow or achillea or millefolii herba).ti,ab.
22. (((a or actaea) adj racemosa) or black cohosh or Black snakeroot or Cimicifuga racemosa).ti,ab.
23. (((a or Aesculus) adj hippocastanum) or horse chestnut or conker tree or Hippocastani semen).ti,ab.
24. (((a or Albizia) adj lebbe#k) or albizia or lebbe#k).ti,ab.
25. (((Allium or a) adj cepa) or onion or Allii cepae bulbus).ti,ab.
26. (((Allium or a) adj sativum) or garlic or Allii sativi bulus).ti,ab.
27. (aloe or Curacao aloes or Barbados aloes or Cape aloes).ti,ab.
28. (((a or Althaea) adj officinalis) or Marshmallow or marsh mallow or Althaeae radix).ti,ab.
29. (((Andrographis or a) adj paniculata) or andrographis).ti,ab.
30. (Angelica or archangelica).ti,ab.
31. (((a or apium) adj graveolens) or celery).ti,ab.
32. (((a or Arctium) adj lappa) or Burdock).ti,ab.
33. (((Arctostaphylos or a) adj uva ursi) or Bearberry or uva ursi or uvae ursi).ti,ab.
34. (((Armoracia or a) adj rusticana) or Horseradish).ti,ab.
35. (((a or Artemisia) adj absinthium) or Wormwood).ti,ab.
36. (((Astragalus or a) adj propinquus) or ((Astragalus or a) adj (membranace?us or membranac*)) or Astragalus or milkvetch or milk vetch).ti,ab.
37. (((Avena or a) adj sativa) or oats or Avenae fructus).ti,ab.
38. (((b or Bacopa) adj monnier#) or Bacopa or brahmi or water hyssop).ti,ab.
39. (((b or Berberis) adj vulgaris) or Barberry).ti,ab.
40. (((b or Boswellia) adj serrata) or Boswellia or frankincen#e).ti,ab.
41. (((B or Bupleurum) adj falcatum) or Bupleurum).ti,ab.
42. (((c or Calendula) adj officinalis) or (Calendula or marigold)).ti,ab.

43. (((c or Camellia) adj sinensis) or green tea).ti,ab.
44. (((Capsicum or c) adj (minimum or annuum or frutescens)) or cayenne or red pepper or bell pepper or hot pepper or chilli or capsicum).ti,ab.
45. (((Cassia or c) adj (angustifolia or senna)) or ((senna or s) adj alexandria) or indian senna).ti,ab.
46. (((c or Centella) adj asiatica) or Gotu kola or pennywort).ti,ab.
47. (((c or Chelidonium) adj majus) or Celandine).ti,ab.
48. (((c or cinnamomum) adj (zeylanicum or cassia or verum or aromaticum)) or cinnamon or Cinnamomi cortex).ti,ab.
49. (((c or Coleus) adj forskohlii) or ((Plectranthus or p) adj barbatus) or Coleus or Forskohlii).ti,ab.
50. (((Commiphora or c) adj (myrrha or molmol)) or myrrh).ti,ab.
51. (((Crataegus or c) adj (oxyacantha or monogyna)) or hawthorn).ti,ab.
52. (((Crocus or c) adj sativus) or saffron).ti,ab.
53. (((Curcuma or c) adj longa) or turmeric or curcumin).ti,ab.
54. (((Cynara or c) adj scolymus) or artichoke).ti,ab.
55. (((Dioscorea or d) adj villosa) or wild yam).ti,ab.
56. (((Drosera or d) adj (rotundifolia or angelica or intermedia)) or sundew).ti,ab.
57. (Echinaceae or Echinacea).ti,ab.
58. (((e or Eleutherococcus) adj senticosus) or Siberian Ginseng or Acanthopanax senticosus).ti,ab.
59. (((elymus or elytr#gia or e or Agropyron or a) adj repens) or couch grass).ti,ab.

60. (((e or Equisetum) adj arvense) or horsetail).ti,ab.
61. (((Eschschol?zia or e) adj californica) or California poppy).ti,ab.
62. ((Eucalyptus or e) adj (globulus or eucalyptus)).ti,ab.
63. (((Eupatorium or e) adj perfoliatum) or Boneset).ti,ab.
64. (((Euphorbia or e) adj hirta) or asthma adj (weed or plant)).ti,ab.
65. (((Euphrasia or e) adj officinalis or rostkoviana) or eyebright).ti,ab.
66. (((Filipendula or f) adj ulmaria) or meadowsweet).ti,ab.
67. (((Frangula or f or rhamnus or r) adj purshiana) or cascara).ti,ab.
68. (((Fucus or f) adj vesiculosus) or bladderwrack).ti,ab.
69. (((Galega or g) adj officinalis) or Goat's rue or galega or french lilac).ti,ab.
70. (((Galium or g) adj aparine) or cleavers).ti,ab.
71. (((Gentiana or g) adj lutea) or gentian or Gentianae radix).ti,ab.
72. (((Geranium or g) adj maculatum) or Cranesbill or geranium).ti,ab.
73. (((Ginkgo or g) adj biloba) or ginkgo or gingko).ti,ab.
74. (((Glycyrrhiza or g) adj glabra) or licorice or Liquiritiae radix or liquorice).ti,ab.
75. (((Gymnema or g) adj sylvestre) or Gymnema).ti,ab.
76. (((Hamamelis or h) adj virginiana) or Witch Hazel).ti,ab.
77. (((Harpagophytum or h) adj procumbens) or Devil's claw).ti,ab.
78. (((Hedera or h) adj helix) or ivy).ti,ab.
79. (((Hemidesmus or h) adj indicus) or Hemidesmus or Indian sarsaparilla).ti,ab.
80. (((Humulus or h) adj lupulus) or hops).ti,ab.
81. (((Hydrastis or h) adj canadensis) or goldenseal).ti,ab.
82. (((Hypericum or h) adj perforatum) or st johns wort).ti,ab.
83. (((Inula or i) adj helenium) or Elecampane).ti,ab.
84. (((Iris or i) adj versicolor) or blue flag).ti,ab.
85. (((Lavandula or l) adj (officinalis or angustifolia or spica or vera)) or Lavender).ti,ab.
86. (((Leonurus or l) adj cardiaca) or Motherwort).ti,ab.
87. (((Linum or l) adj usitatissimum) or Linseed or flaxseed or flax).ti,ab.
88. (((Lycopus or l) adj virginicus) or Bugleweed or Gypsyweed).ti,ab.
89. (((Marrubium or m) adj vulgare) or White Horehound).ti,ab.
90. (((Matricaria or m) adj (chamomilla or recrutita or recruitica)) or C?amomile or Matricariae flos).ti,ab.
91. (((Melaleuca or m) adj alternifolia) or tea tree or Melaleucae aetheroleum).ti,ab.
92. (((Melissa or m) adj officinalis) or Lemon balm or Melissae folium).ti,ab.
93. (Mentha x piperita or peppermint or Mentha balsamea or Menthae piperitae).ti,ab.
94. (((Nigella or n) adj sativa) or black cumin).ti,ab.
95. (((Olea or o) adj europaea) or olive).ti,ab.
96. (((Paeonia or p) adj (officinalis or suffruticosa)) or peony).ti,ab.
97. (((Panax or p) adj (ginseng or notoginseng)) or ginseng).ti,ab.
98. (((Passiflora or p) adj incarnata) or passionflower or passion flower).ti,ab.
99. (((Phytolacca or p) adj (decandra or americana)) or poke root).ti,ab.
100. (((Pimpinella or p) adj anisum) or Aniseed or Anise).ti,ab.
101. (((Piper or p) adj methysticum) or kava).ti,ab.
102. (((Piscidia or p) adj erythrina) or Jamaican dogwood).ti,ab.
103. (((Plantago or p) adj lanceolata) or ribwort).ti,ab.
104. (((Plantago or p) adj ovata) or Psyllium).ti,ab.
105. (((Polygonum or p) adj aviculare) or knotweed).ti,ab.
106. (((Prunus or p) adj serotina) or Wild cherry).ti,ab.
107. (((Ptychopetalum or p) adj olacoides) or Muira puama or Potency wood).ti,ab.
108. (((Rehmannia or r) adj glutinosa) or Rehmannia).ti,ab.
109. (((Rhodiola or r) adj rosea) or Rhodiola or Rhodiolae roseae or rose root or sedum roseum).ti,ab.
110. (((Rosmarinus or r) adj officinalis) or Salvia Rosmarinus or rosemary).ti,ab.
111. (((Rubus or r) adj idaeus) or raspberry or Rubus strigosus).ti,ab.

112. (((Rumex or r) adj crispus) or (yellow or curly) adj dock).ti,ab.
113. (((Salix or s) adj alba) or white willow).ti,ab.
114. (((Salvia or s) adj officinalis) or sage).ti,ab.
115. (((Sambucus or s) adj nigra) or (elder and flower)).ti,ab.
116. (((Schi#andra or s) adj chinensis) or Schi#andra).ti,ab.
117. (((Scutellaria or s) adj baicalensis) or Baikal S#ullcap).ti,ab.
118. (((Scutellaria or s) adj lateriflora) or s#ullcap).ti,ab.
119. (((Serenoa or s) adj (serrulata or repens)) or Saw Palmetto).ti,ab.
120. (((Silybum or s) adj marianum) or St Mary?s Thistle or milk thistle).ti,ab.
121. (((Solidago or s) adj virgaurea) or Goldenrod or Solidago decurrens or Solidaginis virgaureae herba).ti,ab.
122. (((Stellaria or s) adj media) or Chickweed).ti,ab.
123. (((Tanacetum or t) adj parthenium) or Feverfew
124. (((Taraxacum or t) adj officinal*) or Dandelion).ti,ab.
125. (((Thuja or t) adj occidentalis) or Thuja).ti,ab.
126. (((Thymus or t) adj vulgaris) or thyme).ti,ab.
127. (Tilia or (lime flower?) or linden).ti,ab.
128. (((Tribulus or t) adj terrestris) or Tribulus).ti,ab.
129. (((Trifolium or t) adj pratense) or Red clover).ti,ab.
130. (((Trigonella or t) adj foenum graecum) or fenugreek).ti,ab.
131. (((Turnera or t) adj diffusa) or Damiana).ti,ab.
132. (((Ulmus or u) adj (rubra or fulva)) or Slippery elm).ti,ab.
133. (((Urtica or u) adj dioica) or Nettle or (Urticae adj (herba or folium or radix))).ti,ab.
134. (((Vaccinium or v) adj macrocarpon) or Cranberry).ti,ab.
135. (((Vaccinium or v) adj myrtillus) or Bilberry).ti,ab.
136. (((Valeriana or v) adj officinalis) or Valerian).ti,ab.
137. (((Verbascum or v) adj thapsus) or Mullein).ti,ab.
138. (((Verbena or v) adj officinalis) or Vervain).ti,ab.
139. (((Viburnum or v) adj opulus) or Cramp bark).ti,ab.
140. (((Vitex or v) adj agnus castus) or Chaste tree or chasteberry or agnus castus).ti,ab.
141. (((Withania or w) adj somnifera) or Withania or ashwaganda).ti,ab.
142. (((Zanthoxylum or z) adj (clava hercul#s or americanum)) or Prickly ash).ti,ab.
143. (((Zea or z) adj mays) or (corn and silk)).ti,ab.
144. (((Zingiber or z) adj officinal*) or Ginger).ti,ab.
145. (((Ziz#phus or z) adj (jujuba or spinosa)) or Chinese date or jujuba or jujube).ti,ab.
146. or/21-145
147. 1 and 146
148. 147 not 5
149. 20 or 148

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

Ovid syntax

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

CINAHL syntax

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

PubMed syntax

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

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

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

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

Appendix C – Screening criteria

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

1. Duplicate citation
2. Nonhuman study
3. Intervention out of scope (not an in-scope Western herbal medicine)
    a. single herb not on List A of core herbal medicines used by the NHAA[2]
    b. combination herb does not meet eligibility criteria [3]
    c. other (e.g. nutraceuticals, pharmaceutical, other dietary intervention)
4. Population out of scope (healthy participants seeking general wellness)
5. Comparator out of scope (review compares WHM with another WHM)
6. Outcome out of scope (patient experiences/preferences, safety, quality and economic)
7. Publication type out of scope
    a. opinion piece/editorial/commentary
    b. not an intervention study examining effectiveness
8. Study design out of scope (specify)
    a. Non-systematic review, Guideline or HTA assessment
    b. SR of NRSIs or case series
    c. Randomised controlled trial (RCT)
    d. Nonrandomised comparative study
    e. Case series or other
9. Duplicate citation submitted to the Department
    a. SR already identified for this Overview
    b. RCT already included in an eligible SR
10. Publication not available in English [a]
11. Other (specify):
    a. duplicate data (multiple reports arising from the same study)
    b. superseded (SR has been updated or more recent SR is available)
    c. withdrawn
    d. erratum
12. Relevant but additional followup needed (specify) [b]
    a. conference proceeding (data incomplete)
    b. SR protocol only (results not available)
    c. no outcome of interest reported

a. Screening of articles not published in English will be conducted as described in the Section *3.3.1* Reviews published in languages other than English*'*.

b. Articles tagged as relevant but additional followup needed are included but will not be incorporated in the evidence appraisal. These studies may be listed as '*Studies awaiting classification'*, '*Ongoing*', or may be considered when developing conclusions about the '*Overall completeness and applicability of evidence'*.

[2] i.e. herb from other tradition, dosages not permitted in Australia, administration route not in-scope

[3] i.e. does not include at least one of the NHAA core herbal medicines, includes ingredients not on TGA's list of permissible ingredients, includes herbs from other traditions, includes herbal ingredients in dosages not permitted in Australia, administration routes not in-scope, includes other non-herbal ingredients

Appendix D – Risk of bias forms

AMSTAR-2 (systematic reviews)

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

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

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

Source: Shea 2017 (21)

Appendix E – Data extraction forms

Characteristics of included reviews

Review ID	**Author date**				
Review Title					
Review objective					
Author affiliations					
Source of funds					
Declared interests of the review authors					
Review method of analysis	e.g. Narrative, meta-analysis, Guidelines, HTA report				
Inclusion criteria					
Study design					
Population					
Intervention					
Comparator					
Other					
Exclusion criteria					
Study design					
Population					
Intervention					
Comparator					
Other					
Date of documented search (month/year)					
Databases searched (list)	PubMed	Cochrane	Embase	etc.	
Was a non-English database searched?	Yes No Not specified	Other comments			
Were studies in a language other than English included?	Yes No Not specified	Studies were identified through a search of CKNI. Studies reported in non-English language journals were translated before assessment. (p11)			
Outcomes of SR (list)	*(description, timing, measurement tool, other notable features)*				
1	Primary				
2	Secondary				
3	Not specified				
Risk of bias of the included studies as reported in the SR	*(tool used, authors summary)*				
RCTs	Tool used	Authors summary			
NRSIs	Tool used	Authors summary			

Review ID	**Author date**				
Characteristics of studies included in the SR	*(study ID, study design features, setting, other notable features)*				
1	Study ID				
2	Study ID				
add rows as necessary	Study ID				
Authors conclusions (key message)					

Studies meeting the inclusion criteria for this Overview	(*Study ID, no. of participants, Summary RoB, Comments)*			
1	Arab 2016	493	Low risk of bias for all key domains	e.g. Outcome specific details to be added
2				
Add rows as necessary				
Studies excluded from this Overview	*(Study ID, no. of participants, Reasons)*			
1				
2				
Add rows as necessary				

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

Abbreviations:
Notes:

Characteristics of reviews awaiting classification

Study ID	Author date
Study description	Title or other descriptive text
Study design	SR with meta-analysis
Participants	
Intervention	
Comparator	
Outcomes	
Notes	Article written in Korean with no English abstract - awaiting translation

Outcome data from included reviews (as reported by the SR authors)

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

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

Notes:

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

b. Data reported by the SR authors (if available)