



HOMEOPATHY FOR PREVENTING AND TREATING HEALTH CONDITIONS

TECHNICAL REPORT
APPENDICES D TO H

prepared by
HTANALYSTS

for
National Health and Medical
Research Council

NHMRC | Natural Therapies Working
Committee
Canberra ACT 2601

AUGUST 2024

Report information

Authors

Maurin R¹, Rogers E¹, Sullivan E¹, Nolan K¹, Denham I¹, Jorgensen MA¹

¹ HTANALYSTS, Level 8, 46 Kippax Street, Surry Hills NSW 2010 Australia

Dates

This technical report and accompanying evidence evaluation report received approval from the National Health and Medical Research Council (NHMRC) Natural Therapies Working Committee (NTWC) on 16 AUG 2024.

The protocol for the evidence evaluation received approval from the NHMRC NTWC on 13 July 2022 (PROSPERO: CRD42022346433).

History

The National Health and Medical Research Council (NHMRC) has been engaged by the Department of Health and Aged Care (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 natural therapies to be reviewed are Alexander technique, aromatherapy, Bowen therapy, Buteyko, Feldenkrais, homeopathy, iridology, kinesiology, naturopathy, Pilates, reflexology, Rolfing, shiatsu, tai chi, western herbal medicine and yoga. These therapies are amongst those excluded from the private health insurance rebate as of 1 April 2019.

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

This technical report has been developed by HTANALYSTS in conjunction with NHMRC, NTWC, and NTREAP. It provides the appendices (Appendix A to Appendix G) and supplementary data related to an evidence valuation of the effect of homeopathy for preventing and treating health conditions. The main body of evidence is presented in the evidence evaluation report. All associated materials have been developed in a robust and transparent manner in accordance with relevant best practice standards (2-5).

Contents

Report information	ii
List of tables.....	v
List of figures	vii
List of abbreviations.....	ix
Appendix D Details of included studies	1
D1 Diseases of the immune system	1
D1.1 Atopic conditions (allergic rhinitis).....	1
D1.2 Atopic conditions (dermatitis and eczema).....	7
D1.3 Recurrent infections (childhood otitis media)	12
D1.4 Recurrent infections (upper respiratory tract infections)	17
D1.5 Recurrent infections (genitourinary)	24
D2 Mental, behavioural or neurodevelopmental disorders.....	27
D2.1 Anxiety.....	27
D2.2 Depression.....	32
D2.3 Neurodevelopmental disorders.....	37
D3 Sleep-wake disorders	43
D3.1 Insomnia and sleep problems.....	43
D4 Diseases of the nervous system.....	48
D4.1 Headache disorders	48
D5 Diseases of the respiratory system	52
D5.1 Asthma	52
D6 Diseases of the digestive system	59
D6.1 Diarrhoea	59
D6.2 Infantile colic and other digestive disorders.....	63
D6.3 Irritable bowel syndrome.....	67
D7 Diseases of the skin.....	70
D7.1 Psoriasis	70
D8 Diseases of the musculoskeletal system or connective tissue.....	74
D8.1 Arthropathies.....	74
D8.2 Back and/or neck pain	80
D9 Diseases of the genitourinary system.....	85
D9.1 Menopausal symptoms or complaints.....	85
D9.2 Menstrual disorders.....	91
D10 Symptoms, signs or clinical findings, not elsewhere classified.....	97
D10.1 Fatigue conditions	97
D10.2 Fibromyalgia.....	100
Appendix E Risk of bias forms.....	106

Appendix F Characteristics of included studies107

 F1 Study details 107

 F2 Supplementary outcome data 107

Appendix G Differences between protocol & review..... 108

 G1 Methods not implemented..... 108

 G2 Changes from protocol 108

Appendix H How comments from methodological review were addressed..... 110

References.....111

List of tables

Table D-1	Overview of PICO criteria of included studies: Allergic rhinitis.....	1
Table D-2	Outcomes considered by the NTWC to be critical or important for decision-making: Allergic rhinitis.....	4
Table D-3	Overview of PICO criteria of included studies: Dermatitis and eczema.....	7
Table D-4	Outcomes considered by the NTWC to be critical or important for decision-making: Dermatitis and eczema	9
Table D-5	Overview of PICO criteria of included studies: Recurrent otitis media.....	12
Table D-6	Outcomes considered by the NTWC to be critical or important for decision-making: Recurrent otitis media	14
Table D-7	Overview of PICO criteria of included studies: Recurrent upper respiratory tract infections	17
Table D-8	Outcomes considered by the NTWC to be critical or important for decision-making: Recurrent upper respiratory tract infections.....	19
Table D-9	Overview of PICO criteria of included studies: Recurrent urinary tract infections	24
Table D-10	Outcomes considered by the NTWC to be critical or important for decision-making: Recurrent urinary tract infections.....	25
Table D-11	Overview of PICO criteria of included studies: Anxiety.....	27
Table D-12	Outcomes considered by the NTWC to be critical or important for decision-making: Anxiety	29
Table D-13	Overview of PICO criteria of included studies: Depression.....	32
Table D-14	Outcomes considered by the NTWC to be critical or important for decision-making: Depression	34
Table D-15	Overview of PICO criteria of included studies: Neurodevelopmental disorders	37
Table D-16	Outcomes considered by the NTWC to be critical or important for decision-making: Neurodevelopmental disorders.....	40
Table D-17	Overview of PICO criteria of included studies: Insomnia and sleep problems.....	43
Table D-18	Outcomes considered by the NTWC to be critical or important for decision-making: Insomnia and sleep problems	45
Table D-19	Overview of PICO criteria of included studies: Headache disorders	48
Table D-20	Outcomes considered by the NTWC to be critical or important for decision-making: Headache disorders.....	50
Table D-21	Overview of PICO criteria of included studies: Asthma.....	52
Table D-22	Outcomes considered by the NTWC to be critical or important for decision-making: Asthma	54
Table D-23	Overview of PICO criteria of included studies: Diarrhoea.....	59
Table D-24	Outcomes considered by the NTWC to be critical or important for decision-making: Diarrhoea.....	61

Table D-25	Overview of PICO criteria of included studies: Infantile colic and other digestive disorders	63
Table D-26	Outcomes considered by the NTWC to be critical or important for decision-making: Infantile colic and other digestive disorders	65
Table D-27	Overview of PICO criteria of included studies: Irritable bowel syndrome.....	67
Table D-28	Outcomes considered by the NTWC to be critical or important for decision-making: Irritable bowel syndrome	68
Table D-29	Overview of PICO criteria of included studies: Psoriasis.....	70
Table D-30	Outcomes considered by the NTWC to be critical or important for decision-making: Psoriasis.....	72
Table D-31	Overview of PICO criteria of included studies: Arthropathies.....	74
Table D-32	Outcomes considered by the NTWC to be critical or important for decision-making: Arthropathies	76
Table D-33	Overview of PICO criteria of included studies: Back and/or neck pain.....	80
Table D-34	Outcomes considered by the NTWC to be critical or important for decision-making: Back and/or neck pain	81
Table D-35	Overview of PICO criteria of included studies: Menopausal symptoms or complaints.....	85
Table D-36	Outcomes considered by the NTWC to be critical or important for decision-making: Menopausal symptoms or complaints.....	87
Table D-37	Overview of PICO criteria of included studies: Menstrual disorders.....	91
Table D-38	Outcomes considered by the NTWC to be critical or important for decision-making: Menstrual disorders	93
Table D-39	Overview of PICO criteria of included studies: Fatigue conditions	97
Table D-40	Outcomes considered by the NTWC to be critical or important for decision-making: Fatigue conditions.....	98
Table D-41	Overview of PICO criteria of included studies: Fibromyalgia	100
Table D-42	Outcomes considered by the NTWC to be critical or important for decision-making: Fibromyalgia.....	102

List of figures

Figure D-1	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – allergic rhinitis	4
Figure D-2	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Atopic conditions (dermatitis and eczema)	9
Figure D-3	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – otitis media	14
Figure D-4	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Recurrent upper respiratory tract infections	19
Figure D-5	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – recurrent urinary tract infections	25
Figure D-6	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – anxiety	29
Figure D-7	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – depression	33
Figure D-8	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Neurodevelopmental disorders	39
Figure D-9	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Insomnia and sleep problems	44
Figure D-10	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Headache disorders	49
Figure D-11	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – asthma	54
Figure D-12	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – diarrhoea	61
Figure D-13	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – infantile colic and other digestive disorders	64
Figure D-14	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – irritable bowel syndrome	68
Figure D-15	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – psoriasis	71
Figure D-16	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – arthropathies	76
Figure D-17	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Back and/or neck pain	81
Figure D-18	Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Menopausal symptoms or complaints	87

Figure D-19 Risk of bias summary: review authors' judgements about each risk of bias item
expressed as percentages across all RCTs – Menstrual disorders..... 93

Figure D-20 Risk of bias summary: review authors' judgements about each risk of bias item
expressed as percentages across all RCTs – fatigue conditions98

Figure D-21 Risk of bias summary: review authors' judgements about each risk of bias item
expressed as percentages across all RCTs – Fibromyalgia..... 101

List of abbreviations

BRISA	Regional Base of Health Technology Assessment Reports of the Americas
CINAHL	Cumulative Index to Nursing and Allied Health Literature
COMET	Core Outcome Measures in Effectiveness Trials
GRADE	Grading of Recommendations Assessment, Development and Evaluation
ITT	Intent-to-treat
NHMRC	National Health and Medical Research Council
NRSI	Nonrandomised study of an intervention
N TREAP	Natural Therapies Review Expert Advisory Panel
NTWC	Natural Therapies Working Committee
OR	Odds ratios
PAHO	Pan American Health Organization
PICO	Population, Intervention, Comparator, Outcome
PP	Per protocol
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised controlled trial
RoB	Risk of bias
RR	Risk ratios
SR	Systematic review
SD	Standard deviation
TIDIER	Template for Intervention Description and Replication

Appendix D Details of included studies

This appendix documents the studies that met the prespecified inclusion criteria for a systematic review on the effect of homeopathy for preventing and treating any health condition and were prioritised at the population prioritisation phase. It provides an overview of the PICO criteria of these studies, a summary of the risk of bias assessment, and results of the data synthesis for the main comparison.

Additional details concerning the risk of bias judgements for each study are provided in **Appendix E** and characteristics of the included studies are provided in **Appendix F1**. Outcome data for outcomes considered to be critical or important for this review are provided in **Appendix F2**.

D1 Diseases of the immune system

D1.1 Atopic conditions (allergic rhinitis)

D1.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-1. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-1 Overview of PICO criteria of included studies: Allergic rhinitis

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Aabel 2000a (6)	quasi RCT	Hay fever (allergic rhinitis) (birch pollen)	Non-individualised, oral (Betula 30c)	Placebo	None reported	Symptom severity; use of rescue medication
Aabel 2000b (7)	RCT	Hay fever (allergic rhinitis) (birch pollen)	Non-individualised, oral (Betula 30c)	Placebo	None reported	Symptom severity; use of rescue medication
Aabel 2001 (8)	RCT	Hay fever (allergic rhinitis) (birch pollen)	Non-individualised, oral (Betula 30c)	Placebo	None reported	Symptom severity; use of rescue medication
Kim 2005 (9)	RCT	Hay fever (allergic rhinitis)	Non-individualised, sublingual combination	Placebo	None reported	Symptom severity; work impairment; health-related quality of life
Liu 2013 (10)	RCT	Hay fever (allergic rhinitis)	Individualised, sublingual combination (selected according to the common symptoms of allergic rhinitis)	Placebo	None reported	Symptom severity; inflammatory biomarkers

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Naidoo 2013 (11)	RCT	Hay fever (allergic rhinitis) (cat fur)	Non-individualised, oral combination (Cat saliva 9CH and Histaminum 9CH)	Placebo	None reported	Objective allergic response
Reilly 1984 (12, 13)	RCT	Hay fever (allergic rhinitis)	Non-individualised, oral (30c mixed grass pollen)	Placebo	Escape antihistamine (if needed)	Symptom severity
Taylor 2000 (14-16)	RCT	Hay fever (allergic rhinitis)	Non-individualised, oral (30c of principal inhalant allergen)	Placebo	None reported	Symptom severity; use of rescue medication
Wiesenauer 1995 (17)	quasi RCT	Hay fever (allergic rhinitis)	Non-individualised, oral (Galphimia glauca D4)	Placebo	None reported	Symptom severity; use of additional medication; compliance; satisfaction
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus 'other' intervention**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D1.1.2 Risk of bias per item

The risk of bias for each item in the included studies for allergic rhinitis is described below and shown graphically in Figure D-1 (details are provided in Appendix E).

Bias arising from the randomisation process

Four studies (Aabel 2000a, Aabel 2000b, Liu 2013, Reilly 1984) were assessed at low risk of bias for this domain, participant allocation was randomised, and information was provided for the method used to generate the randomisation sequence and method of allocation concealment. Five studies (Aabel 2001, Kim 2005, Naidoo 2013, Taylor 2000, Wiesenauer 1995) were assessed to have some concerns for this domain. Concerns arose due to lack of reporting on allocation concealment (Kim 2005, Taylor 2000) and lack of reporting on baseline characteristics (Aabel 2001, Naidoo 2013). In Wiesenauer 1995, participants were allocated to placebo or intervention via stratified randomisation, meaning randomisation was done within each stratum; however, strata were defined by the treating physician and no details of the strata definition were provided.

Bias due to deviations from intended interventions

Seven studies (Aabel 2000a, Aabel 2000b, Aabel 2001, Kim 2005, Liu 2013, Naidoo 2013, Taylor 2000) were assessed at low risk of bias for this domain as participants and study staff were blinded, and analysis was by intent-to-treat.

Two studies (Reilly 1984, Wiesenauer 1995) were assessed at high risk of bias for this domain. In Reilly 1984, the number of outcomes responses exceeded the number of participants randomised and it was unclear how many randomised participants were analysed. In Wiesenauer 1995, a per protocol analysis was used, excluding participants who used other medicines. The analysis excluded a total of 32 participants.

Bias due to missing outcome data

Four studies (Aabel 2000a, Aabel 2001, Naidoo 2013, Taylor 2000) were assessed at low risk of bias for this domain as data were available for all, or nearly all participants randomised in the trial. Two studies (Aabel 2000b, Reilly 1984) were assessed to have some concerns for this domain as there was some missing data with no analysis presented to adjust for missingness. However, it was not considered likely that missingness was related to the true value of the outcome.

Three studies (Kim 2005, Liu 2013, Wiesenauer 1995) were assessed at high risk of bias for this domain, as all studies had missing data and provided no analysis for the missing data. In one study (Kim 2005) patients dropped out of the study, including 2 (10%) in the homeopathic group and 4 (20%) in the placebo group, authors note that the discontinuation was likely due to lack of response to treatment, as such it was considered likely that the missingness in the outcome will depend on the true value. In Liu 2013 it is plausible that withdrawals in treatment phase were related to ineffective treatment, as the rate of missingness was unbalanced between arms. In Wiesenauer 1995 authors note that reasons for patient withdrawal include incomplete documentation, self-medication or additional hay fever medication administered by the physician, likely due to perceived effectiveness of the allocated treatment.

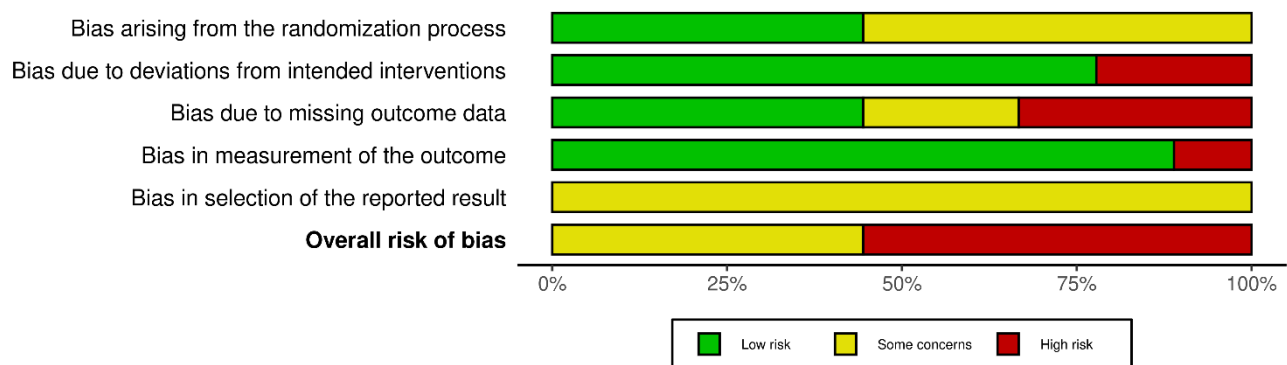
Bias in measurement of the outcome

Eight studies (Kim 2005, Lie 2013, Reilly 1984, Taylor 2000, Aabel 2000a, Aabel 2000b, Naidoo 2013, Wiesenauer 1995) were judged at low risk of bias for this domain as the outcomes were measured appropriately, all studies were double blind and outcome measurement was consistent between groups. One study (Aabel 2001) was assessed at high risk of bias for this domain. Due to low pollen counts during the trial period, participants in Aabel 2001 were asked to continue recording outcomes for as long as possible, leading to variable reporting duration between participants and groups.

Bias in selection of the reported result

All 9 studies (Kim 2005, Lie 2013, Reilly 1984, Taylor 2000, Aabel 2000a, Aabel 2000b, Aabel 2001, Naidoo 2013, Wiesenauer 1995) were assessed to have some risk of bias for this domain, as no pre-specified analysis plan was published for comparison to the final analysis. There was no evidence that the selection of results was based on multiple eligible outcome measurements within the outcome domain.

Figure D-1 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – allergic rhinitis



D1.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with allergic rhinitis are listed in Table D-2.

Table D-2 Outcomes considered by the NTWC to be critical or important for decision-making: Allergic rhinitis

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison ?	Aabel 2000a	Aabel 2000b	Aabel 2001	Kim 2005	Liu 2013	Naidoo 2013	Reilly 1984	Taylor 2000	Wiesenauer 1995
Symptom severity	Total Nasal Symptom Score, VAS (or other validated measure)	Critical	Yes	✓^	✓^	✓^	--	✓	--	✓	✓	✓^
HRQoL	Rhinoconjunctivitis QLQ (or other validated measure)	Critical	Yes	--	--	--	✓	--	--	--	--	--
Medication use	Patient diary	Critical	Yes	✓^	✓^	✓^	--	--	--	✓	✓^	--

Abbreviations: HRQoL, health-related quality of life; QLQ, quality of life questionnaire; VAS, visual analogue scale

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

^ Study data was not able to be extracted or was not able to be included in the data synthesis

D1.1.3.1 Primary Comparison (vs placebo)

Nine studies comparing homeopathy with placebo in people with allergic rhinitis were eligible for this comparison and contributed data to 3 outcomes considered critical or important for this review.

There were 8 ongoing studies (973 participants) and 4 studies awaiting classification (525 participants) that could have contributed data to this comparison. Of these, nonreporting of results was suspected in at least 4 studies listed as ongoing (503 participants) because the p-value, magnitude or direction of the results generated were considered unfavourable by the study investigators.

Symptom severity

Two studies (Reilly 1984, Taylor 2000) reported symptom severity measured using a 100 mm visual analogue scale (VAS) at end of treatment (range: 4 to 5 weeks).

The VAS is subjective tool that can be used to measure a variety of outcomes. It is measured on a continuous scale (mm) from 0 (no symptoms) to 100 (greatest symptom severity), with higher scores indicating a higher intensity of symptoms. Both studies reported results as the mean change from baseline to the end of the study period. In Taylor 2000, this corresponds to change from baseline to week 5, while the intervention completed after week 3 (i.e. results correspond to the end of follow up).

Pooled results from 2 studies (total 158 participants) showed no difference in symptom severity between homeopathy and placebo (SMD -0.32 ; 95% CI $-0.70, 0.06$; $p = 0.10$).

One study (Liu 2013) reported symptom severity measured using the Total Nasal Symptom Score (TNSS) at end of treatment (4 weeks).

The TNSS measures nasal symptoms across 4 items (nasal congestion, sneezing, nasal itching, and rhinorrhoea) on a 4-point scale (0-3). The total score is the sum of the scores across each symptom (scores range 0-12), where a higher score indicates more severe symptoms (18). Scores are reported as mean change from baseline to end of treatment. Results from one study (36 participants) show no difference in symptom severity between homeopathy and placebo (SMD -0.22 ; 95% CI $-0.90, 0.46$; $p = 0.53$).

Pooled results across 3 studies (194 participants) suggest an improvement in symptom severity in the homeopathy group compared to placebo (SMD -0.32 ; 95% CI $-0.60, -0.03$; $p = 0.03$; $I^2 = 0\%$). Based on Cohen's guidance, this was considered a small change in favour of homeopathy (i.e. SMD < 0.5).

(GRADE: Very Low)

In a sensitivity analysis examining the impact of 2 RCTs at high risk of bias, the size of the effect estimate decreased and the difference between groups was no longer considered important (i.e. SMD < 0.2) (SMD -0.06 ; 95% CI $-0.62, 0.49$; $p = 0.82$).

Results from 4 additional studies including 310 participants (Aabel 2000a, Aabel 2000b, Aabel 2001, Wiesenauer 1995) were not able to be extracted for meta-analysis. The studies reported symptom severity measured using a VAS or Numeric Rating Scale. In Aabel 2000a, Aabel 2000b and Aabel 2001, results were reported in a graphic format as the rating each day of the trial. Study authors reported no difference between the homeopathy and placebo arms for the majority of study days in these trials. In Aabel 2001 (which included participants from Aabel 2000a and Aabel 2000b), data was reported for each arm as mean and 95% CI in table format. Data extraction from this table was attempted, but there was insufficient detail reported by the trialists to accurately calculate standard deviations. As such, data from Aabel 2001 was not able to be included in the meta-analysis. In Wiesenauer 1995, results were reported as the proportion of people who experienced an improvement in nasal and ocular symptoms. The study authors report no difference in nasal symptoms and a significant difference in favour of homeopathy for ocular symptoms.

Quality of life

One study (Kim 2005) reported quality of life measured with the Rhinoconjunctivitis Quality of Life Questionnaire at end of treatment (4 weeks).

The Rhinoconjunctivitis Quality of Life Questionnaire is a 28-item disease-specific quality of life questionnaire that measures 7 domains: activity limitations, sleep problems, nose symptoms, eye symptoms, non-nose/eye symptoms, practical problems, and emotional function (19). Each item is rated on a 7-point scale (0 to 6), where a higher score indicates a worse quality of life. The final score is calculated as the mean across each domain. The minimally important difference has been reported to be a change of 0.5 points (19). Results from one study (34 participants) show no difference in quality of life between homeopathy and placebo (MD -0.40; 95% CI -1.10, 0.30; $p = 0.26$). (*GRADE: Very Low*)

No sensitivity analysis examining the impact of RCTs at high risk of bias was conducted, as only one RCT contributed data to this analysis.

Medication use

One study (Reilly 1984) reported medication use based on the number of antihistamine tablets taken during the trial period (5 weeks).

Antihistamine tablets are used as rescue medication for people with allergic rhinitis, and act to treat symptoms by reversing the action of histamine. Fewer antihistamine tablets required is better. Results from one study (108 participants) show a reduction in the use of antihistamines in the homeopathy group compared to placebo (MD -8.50; 95% CI -14.67, -2.33; $p = 0.007$). (*GRADE: Low*)

No sensitivity analysis examining the impact of RCTs at high risk of bias was conducted, as only one RCT contributed data to this analysis.

Additionally, 4 RCTs (240 participants) measured medication use but did not report it in a format that could be extracted for meta-analysis (Aabel 2000a, Aabel 2000b) or did not report it at all (Aabel 2001, Taylor 2000).

D1.1.3.2 Secondary Comparison (vs 'inactive' control)

There were no studies identified comparing homeopathy with inactive control in people with allergic rhinitis.

D1.1.3.3 Tertiary Comparison (vs other)

There were no studies identified comparing homeopathy with other interventions in people with allergic rhinitis.

D1.2 Atopic conditions (dermatitis and eczema)

D1.2.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-3. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-3 Overview of PICO criteria of included studies: Dermatitis and eczema

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus control (placebo)*						
Carello 2017 (20)	RCT	Atopic dermatitis	Non-individualised homeopathy, oral combination (Galiam-Heel)	Placebo	Standard medical care (topical steroids, antibiotics and antihistamines)	Disease severity, inflammatory biomarkers, itching, sleep disturbances, medication use, treatment safety
Dey 2022 (21)	RCT	Atopic dermatitis	Individualised homeopathy, oral	Placebo	None reported	Disease severity, quality of life, disease burden
Vickers 2000 (22, 23)	RCT	Atopic dermatitis	Individualised homeopathy, oral	Placebo	None reported	Disease severity, sleep, itching, medication use, quality of life
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus 'other' intervention**						
No studies found						

Note: cH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D1.2.2 Risk of bias per item

The risk of bias for each item in the included studies for atopic dermatitis and eczema is described below and shown graphically in Figure D-2 (details are provided in Appendix E).

Bias arising from the randomisation process

Two studies (Dey 2022, Vickers 2000) were assessed at low risk of bias for this domain as they provided information regarding the generation of the randomisation sequence and method of allocation concealment. There were significant differences in age and socioeconomic status in Dey 2002, however this was not considered likely to be due to issues with randomisation.

One study (Carello 2017) was assessed to have some concerns for this domain, as the method of allocation concealment was unclear. It was stated that a randomisation list was computer generated, but participants were allocated to groups according to sequential order of enrolment. It was considered plausible that allocation was not concealed from the enrolling investigator. There were no significant differences between groups at baseline.

Bias due to deviations from intended interventions

Three studies (Carello 2017, Dey 2022, Vickers 2000) were assessed at low risk of bias for this domain as the participants and research staff were both blinded to the intervention, and an appropriate analysis was undertaken.

Participants in Vickers 2000 were allocated into 4 groups, including one group where participants received unblinded homeopathic medication, as such they were aware of their assigned arm during the trial. A differential rate of drop out was observed between the blinded and unblinded treatment arms, considered likely to be due to the trial context. This was not balanced between arms. However, as only the blinded results are reported in the analysis, an overall low risk of bias was considered for this domain.

Bias due to missing outcome data

Two studies (Carello 2017, Dey 2022) were assessed to have some concerns for this domain due to the proportion of missing data with no analysis presented to adjust for missingness. The missingness was not considered likely to be due to the true value of the outcome.

One study (Vickers 2000) was assessed at high risk of bias for this domain due to the proportion of missing outcome data, which could plausibly have been related to the outcome. The study authors report a differential rate of missingness between the blinded and unblinded treatment arms, suggesting that knowledge of treatment influenced discontinuation.

Bias in measurement of the outcome

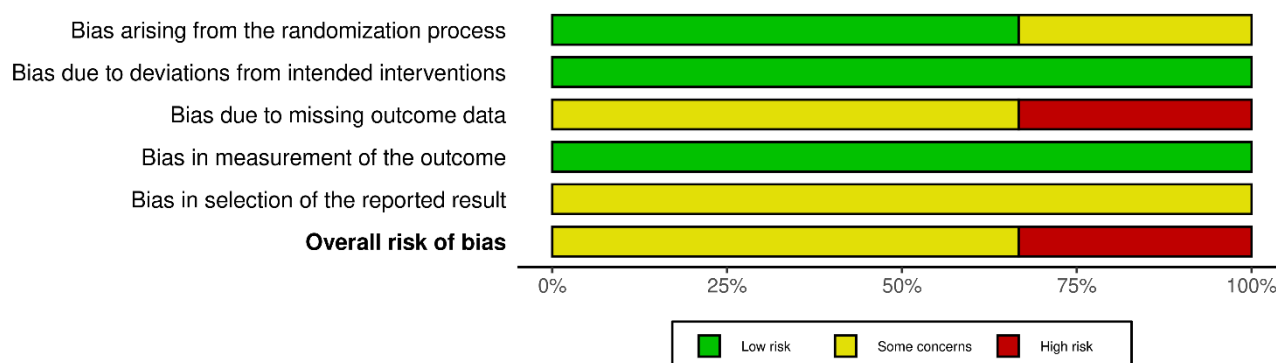
Three studies (Carello 2017, Dey 2022, Vickers 2000) were judged at low risk of bias for this domain as the outcomes were measured appropriately, and ascertainment of the outcome was consistent between groups.

Participants in Vickers 2000 were allocated into 4 groups, including one group where patients received unblinded homeopathic medication. The study included patient reported outcomes, and as the participants were not blinded to the intervention they received, outcome assessors were aware of the intervention received, and it was considered likely that assessment of the outcome was influenced by knowledge of the outcome received. The study authors of Vickers 2000 note that patients who were blind to the intervention were more likely to drop out of the trial than those participants who received the unblinded intervention, resulting in unblinded participants contributing more data to the trial. However, as only the blinded results are reported in the analysis, an overall low risk of bias was considered for this domain.

Bias in selection of the reported result

Three studies (Carello 2017, Dey 2022, Vickers 2000) were judged to have some concerns for this domain as no pre-specified analysis plan was available, however there was no evidence that the selection of results was based on multiple eligible outcome measurements within the outcome domain.

Figure D-2 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Atopic conditions (dermatitis and eczema)



D1.2.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with dermatitis or eczema are listed in Table D-4.

Table D-4 Outcomes considered by the NTWC to be critical or important for decision-making: Dermatitis and eczema

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Carello 2017	Dey 2022	Vickers 2000
Disease severity	Severity Scoring of Atopic Dermatitis, VAS (or other validated measure)	Critical	Yes	✓	✓	✓
Health-related quality of life	Dermatology Life Quality Index	Critical	Yes	--	✓	✓
Medication use	Patient diary	Critical	Yes	✓	--	✓
Itching	Numeric Rating Scale (or other validated measure)	Critical	Yes	--†	--	--†
Skin condition	Numeric Rating Scale (or other validated measure)	Critical	Yes	--	--	--†

Abbreviations: VAS, visual analogue scale

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

--No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

† No study result is available for inclusion due to insufficient detail provided on the outcome measure. The outcome was measured and reported by the study authors.

D1.2.3.1 Primary Comparison (vs placebo)

Three studies comparing homeopathy with placebo in people with atopic dermatitis were eligible for this comparison and contributed data to 5 outcomes considered critical or important for this review.

There were 2 studies (40 participants) awaiting classification that could have contributed data to 2 outcomes (symptom severity, quality of life) for this comparison.

Disease severity

Three studies (Carello 2017, Dey 2022, Vickers 2000) reported disease severity measured using the Severity Scoring of Atopic Dermatitis (SCORAD) or a 100 mm visual analogue scale at the end of treatment (range: 3 to 8 months).

The SCORAD is a clinical tool designed to assess the extent and severity of atopic dermatitis. It can be completed by a clinician or self-report (Patient-Oriented SCORAD). The SCORAD includes 3 components measuring dermatitis area (A), intensity (B), and subjective symptoms of itch and sleeplessness (C). The total score is calculated based on the scores for each domain to give a total score out of 103, where higher scores are indicative of more severe disease (24). Results from two studies (126 participants) found no difference between homeopathy and placebo in disease severity in people with atopic dermatitis (SMD -0.32 ; 95% CI $-0.67, 0.03$; $p = 0.08$).

The VAS is subjective tool that can be used to measure a variety of outcomes. It is measured on a continuous scale (mm) from 0 (no symptoms) to 100 (greatest symptom severity), with higher scores indicating more severe disease. Results from one study (27 participants) found no difference between homeopathy and placebo in disease severity in people with atopic dermatitis (SMD -0.16 ; 95% CI $-0.92, 0.66$; $p = 0.68$).

Pooled results from 3 studies (153 participants) suggest a reduction in disease severity with homeopathy compared to placebo (SMD -0.29 ; 95% CI $-0.61, 0.03$; $p = 0.08$; $I^2 = 0\%$) (*GRADE: Low*). Based on Cohen's guidance, this would be considered a small change (i.e. SMD <0.5).

In a sensitivity analysis examining the impact of 1 RCT at high risk of bias, the effect size did not meaningfully change (SMD -0.32 ; 95% CI $-0.67, 0.03$; $p = 0.08$).

Quality of life

Two studies (Dey 2022, Vickers 2000) reported quality of life measured with the Dermatology Life Quality Index (DLQI) at end of treatment (3 months).

The DLQI is a disease-specific quality of life measure for dermatology patients. It includes 10 questions across areas such as symptoms, daily activities, leisure, work or school, personal relationships, and the side-effects of treatment (25). Each item is scored on a 4-point scale (0 to 3), with the total score being a sum of each item to give a total score of 0 to 30. Higher scores indicate worse quality of life. Studies assessing the MCID for the DLQI have varied from 3 to 5 points (25).

Results from two studies (87 participants) suggest no difference in health-related quality of life between homeopathy and placebo, with some heterogeneity (MD -0.55 ; 95% CI $-2.02, 0.93$; $p = 0.47$; $I^2 = 66\%$) (*GRADE: Low*).

In a sensitivity analysis examining the impact of 1 RCT at high risk of bias, the effect size did not meaningfully change (MD -1.60 ; 95% CI $-3.40, 0.20$; $p = 0.08$).

Medication use

One study (Vickers 2000) reported medication use measured using a self-reported diary of steroid cream use at end of treatment (12 weeks).

Topical steroids are used in atopic dermatitis to manage and treat eczema flares (26). Vickers 2000 measured steroid cream use using a 5-point Likert scale where a higher score indicated less frequent use. Results from 1 study (27 participants) suggest no difference in medication use between homeopathy and placebo (SMD 0.23 ; $-0.53, 0.99$; $p = 0.56$) (*GRADE: Very Low*).

No sensitivity analysis examining the impact of RCTs at high risk of bias was conducted, as only one study was contributing data.

Results from one additional study (66 participants) were not able to be included in the meta-analysis, as the study reported the proportion of participants using antihistamines in each treatment arm. The study authors report no difference in antihistamine use between homeopathy and placebo.

Itching

Two studies (Carello 2017, Vickers 2000) reported itching using measures which were not defined, including the direction of effect. These results were not able to be interpreted or included in the meta-analysis.

Skin condition

One study (Vickers 2000) reported skin condition using a 10-point digital scale, however the measure was not further defined, including the direction of effect. Results were not able to be interpreted or included in the meta-analysis.

D1.2.3.2 Secondary Comparison (vs 'inactive' control)

There were no studies identified comparing homeopathy with inactive control in people with atopic dermatitis.

D1.2.3.3 Tertiary Comparison (vs other)

There were no studies identified comparing homeopathy with other interventions in people with atopic dermatitis.

D1.3 Recurrent infections (childhood otitis media)

D1.3.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-5. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-5 Overview of PICO criteria of included studies: Recurrent otitis media

STUDY ID	STUDY DESIGN	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Jacobs 2001 (27, 28)	RCT	Children with acute otitis media	Individualised homeopathy, oral	Placebo	Consultations with a homeopath	Symptom severity
Pedrero-Escalas 2016 (29)	RCT	Children with otitis media with effusion	Non-individualised, oral (Agraphis nutans 5CH, Thuya Occidentalis 5CH, Kalium muriaticum 9CH and Arsenicum iodatum)	Placebo	Aerosol therapy with corticosteroids and mucolytics	Infection frequency
Homeopathy versus inactive control*						
Harrison 1999 (30)	Quasi RCT	Children with otitis media with effusion	Individualised homeopathy + monthly consultations with a homeopath	Control (no intervention)	Standard care (access to GPs, specialists and antibiotic treatment as required)	Study did not report any priority outcomes
Taylor 2011 (31)	RCT	Children with acute otitis media	Non-individualised, ear drops (Pulsatilla, Chamomilla, Sulphur, Calcarea carbonica, Belladonna, and Lycopodium, all 30c)	Control (no intervention)	Standard care (antibiotics, paracetamol, ibuprofen or benzocaine ear drops)	Symptom severity
Taylor 2014 (32, 33)	RCT	Children with acute otitis media	Non-individualised, ear drops (Pulsatilla, Chamomilla, Sulphur, Calcarea carbonica, Belladonna, and Lycopodium, all 30c)	Control (no intervention)	Standard care (including use of analgesics and directions on when to fill antibiotic prescriptions)	Symptom severity
Homeopathy versus 'other' intervention**						
Sinha 2012 (34)	RCT	Children with acute otitis media	Individualised homeopathy	Active control (symptomatic relief with analgesics, anti-inflammatory and antipyretics)	All groups received antibiotics if less than a 50% improvement was observed in the first 3 days	Symptom severity

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial.

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D1.3.2 Risk of bias per item

The risk of bias for each item in the included studies for recurrent childhood otitis media is described below and shown graphically in Figure D-3 (details are provided in Appendix E).

Bias arising from the randomisation process

Four studies (Jacobs 2001, Pedrero-Escalas 2016, Sinha 2012, Taylor 2014) were assessed to have low concerns in this domain. These studies described their randomisation processes in sufficient detail and reported either no significant baseline differences or slight differences that were not likely due to the randomisation process. There were some concerns of bias in one study (Taylor 2011) as information about the antibiotic treatment plan for participants was not presented, and it was judged likely to have an impact on the outcomes reported.

One study (Harrison 1999) was at high risk of bias for this domain. The study authors suggested the study may have been compromised by the possibility that randomisation was unconcealed. There were some differences noted in baseline characteristics between the groups, further suggesting an issue with the randomisation process and allocation concealment.

Bias due to deviations from intended interventions

All 6 studies (Harrison 1999, Jacobs 2001, Pedrero-Escalas 2016, Sinha 2012, Taylor 2011 and Taylor 2014) were judged to be at low risk of bias for this domain.

Three studies (Jacobs 2001, Pedrero-Escalas 2016 and Sinha 2012) were placebo-controlled, double-blind trials. The participants and those delivering the interventions were blinded to treatment allocations in 3 studies (Harrison 1999, Taylor 2011 and Taylor 2014). The only deviations reported in these studies was non-completion by some participants, but this was deemed not likely due to the randomisation process.

Bias due to missing outcome data

Four studies (Harrison 1999, Jacobs 2001, Taylor 2011 and Taylor 2014) were deemed to have some concerns in this domain due to the proportion of missing outcome data. Reasons for participant drop-out were either not related to health status, or not provided.

Two studies (Pedrero-Escalas 2016 and Sinha 2012) were deemed to have high concerns in this domain. This was due to the proportion of missing outcome data and reasons for participant drop-out relating to health status.

Bias in measurement of the outcome

Four studies (Harrison 1999, Jacobs 2001, Pedrero-Escalas 2016 and Sinha 2012) were assessed to have low concerns in this domain. In these studies, the outcome assessors were blinded and used validated outcome measures that were consistent across treatment groups.

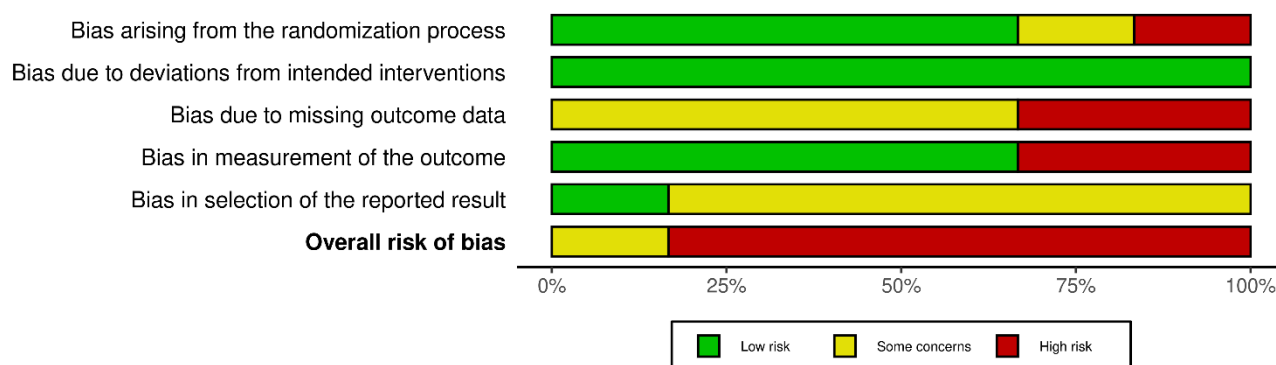
Two studies (Taylor 2011 and Taylor 2014) were assessed to have high concerns in this domain. In these studies, the outcome assessors were not blinded to treatment allocation. It is therefore possible that knowledge of the intervention could have influenced the measurement of the outcome, although there was no evidence to suggest this had occurred.

Bias in selection of the reported result

Five studies (Harrison 1999, Jacobs 2001, Sinha 2012, Taylor 2011 and Taylor 2014) were assessed to have some concerns in this domain. In these studies there was no evidence of multiple eligible outcome measurements or analysis, however no pre-specified analysis plans were available for comparison.

One study (Pedrero-Escalas 2016) was deemed to have low concerns in this domain. In this study there was no evidence of multiple eligible outcome measurements or analysis, and the study protocol was available for comparison.

Figure D-3 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – otitis media



D1.3.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in children with recurrent otitis media are listed in Table D-6.

Table D-6 Outcomes considered by the NTWC to be critical or important for decision-making: Recurrent otitis media

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Harrison 1999	Jacobs 2001	Pedrero-Escalas 2016	Sinha 2012	Taylor 2011	Taylor 2014
Infection frequency	Number of episodes of otitis media	Critical	Yes	--	--	✓	--	--	--
	Number who experience recurrent infection	Critical	Yes	--	--	✓	--	--	--
Infection duration	Number of days per episode (or per year)	Critical	No	--	--	--	--	--	--
Symptom severity	Any validated measure	Critical	Yes	--	✓^	--	✓^	✓^	✓
Quality of life	Any validated measure	Critical	No	--	--	--	--	--	--

Abbreviations: VAS, visual analogue scale.

✓ A study result is available for inclusion in the synthesis.

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators.

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results.

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study.

^ Study data was not able to be extracted.

D1.3.3.1 Primary Comparison (vs placebo)

Two RCTs (Jacobs 2001, Pedrero-Escalas 2016) comparing homeopathy with placebo in children with recurrent otitis media were eligible for this comparison. One study (Pedrero-Escalas 2016) contributed data to 2 outcomes considered critical or important for this review.

There was one ongoing study (target 142 participants) comparing homeopathy with placebo in people with chronic suppurative otitis media that could contribute data to one outcome considered critical or important for this review, but recruitment is not yet complete.

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data for all outcomes.

Infection frequency

One study (total 96 participants) reported infection frequency, measured by the number of participants who experienced an episode of acute otitis media (AOM) during the study period (3 months) and the number who had a recurrent episode of otitis media with effusion (OME) (Pedrero-Escalas 2016). Recurrence was defined as having a positive pneumatic otoscopy (PNO)¹ mid-treatment (second visit), which changed to a negative PNO² at the end of treatment (third visit).

The results showed that there was no difference in infection frequency between the homeopathy group and the placebo group at the end of the study for both outcome measures (*GRADE: Low*):

- number who experienced an episode of AOM: RR 0.78; 95% CI 0.38, 1.57; $p = 0.48$
- number who had a recurrence of OME: RR 0.43; 95% CI 0.09, 2.13; $p = 0.30$.

Symptom severity

One study (total 75 participants) reported symptom severity measured using a daily symptom diary that was parent recorded 3 times daily for the first 3 days (Jacobs 2001). Daily symptom diaries included information on pain, fever, irritability, appetite, energy level, sleep, concurrent upper respiratory tract symptoms and number of doses of study medications given. The diary symptom scores were calculated by assigning numerical values for each of the 7 factors of the symptom diary and totalling, with the maximum possible score being 9 (higher scores indicated greater symptom severity).

Results from one study (Jacobs 2001) (total 75 participants) were not able to be included in the data synthesis as the study did not report means or SD. The authors suggested that symptoms severity decreased faster over time and favoured homeopathy, but the difference between groups was not significantly different at 72-hours ($p > 0.05$) (*GRADE: Very Low*).

D1.3.3.2 Secondary Comparison (vs 'inactive' control)

Two RCTs (Taylor 2011, Taylor 2014) and one quasi RCT (Harrison 1999) comparing homeopathy with an inactive control (no intervention) in children with recurrent otitis media were eligible for this comparison. Two studies (Taylor 2011, Taylor 2014) contributed data to one outcome considered critical or important for this review, but data from one study (Taylor 2011) were incomplete and not able to be included in the meta-analysis. One study (Harrison 1999) did not report any outcomes that were considered critical or important for this review.

There were no ongoing studies and one study awaiting classification (total 90 participants) that was published in a language other than English and could have contributed data to this comparison, but it did not appear to report any outcomes considered critical or important to this review.

¹ presence of tympanic motility

² absence of tympanic motility

Symptom severity

Two studies (Taylor 2011, Taylor 2014) reported symptom severity measured using the ear treatment group symptom questionnaire (ETG-5) at the end of treatment. The ETG-5 is a parent-rated 5-item questionnaire developed to quantify the severity of acute otitis media symptoms in children. The symptoms measured include fever, earache, or tugging, feeding, irritability, and sleep, which have 3 ratings of severity (scored as 0, 4 or 7). The overall ETG-5 score comprises the sum of the individual symptom scores (total ranging from 0 to 35) with higher scores indicating worse symptom severity.

One RCT (Taylor 2011) (total 119 participants) suggested a faster reduction in symptom severity in the homeopathy group compared to the inactive control (no intervention) group, with the difference in ETG-5 scores reported to be significant at assessments number 2 ($p = 0.04$) and 3 ($p = 0.002$). However, the difference was not significant at the end of treatment ($p = 0.36$). The study reported mean scores, but did not report SD, SE or CIs, therefore these results were not able to be included in the meta-analysis.

One RCT (Taylor 2014) (total 210 participants) reported there were no differences in ETG-5 scores between the homeopathy group and the inactive control (no intervention) group at 5 to 7 days, after adjusting for differences in baseline ETG-5 scores (MD 1.30; 95% CI -0.11, 2.71; $p = 0.14$). (*GRADE: Low*).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

D1.3.3.3 Tertiary Comparison (vs other)

One RCT comparing homeopathy with 'other' interventions in people with recurrent childhood otitis media were eligible for this comparison and contributed data for one priority outcome. Data from this study are presented in Appendix F2 Supplementary outcome data.

D1.4 Recurrent infections (upper respiratory tract infections)

D1.4.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-7. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-7 Overview of PICO criteria of included studies: Recurrent upper respiratory tract infections

STUDY ID	STUDY DESIGN	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
de Lange de Klerk 1993 (35-37)	RCT	Children with URTIs	Individualised homeopathy	Placebo	Conventional treatment (as prescribed by the participants GP). All groups received consultations with a homeopath	Infection frequency Medication use
Furuta 2017 (38)	Quasi RCT	Children with recurrent tonsillitis	Individualised homeopathy (30CH) + Non-individualised, oral (Baryta carbonica 6CH) + isopathic 12CH combination (β -Streptococcus, S. aureus, H. influenzae)	Placebo	Participants who developed tonsillitis were treated with antimicrobial agents. All groups received consultations with a homeopath	Infection frequency
Steinsbekk 2004 (39-42)	RCT	Children with URTIs	Individualised homeopathy + consultations with a homeopath OR Self-prescribed individualised homeopathy	Placebo	All groups were allowed treatment of the participant's choice other than homeopathic treatments	Infection frequency Medication use
Homeopathy versus inactive control*						
Palm 2017 (43)	RCT	Children and adults with recurrent tonsillitis	Non-individualised homeopathy, oral complex	Inactive control (no intervention)	Standard care consisting of local antiseptics and local anaesthetics. Antibiotics were given as rescue medication	Infection frequency Symptom severity Medication use
Steinsbekk 2004 (39-42)	RCT	Children with URTIs	Individualised homeopathy OR Self-prescribed individualised homeopathy	Inactive control (no intervention)	All groups were allowed treatment of the participant's choice other than homeopathic treatments	Infection duration Medication use
Homeopathy versus 'other' intervention**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: GP, general practitioner; RCT, randomised controlled trial; URTI, upper respiratory tract infection.

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix E2](#)) if they reported data for outcomes considered critical or important to this review.

D1.4.2 Risk of bias per item

The risk of bias for each item in the included studies for recurrent respiratory tract infections is described below and shown graphically in Figure D-4 (details are provided in Appendix E).

Bias arising from the randomisation process

Three studies (de Lange de Klerk 1993, Palm 2017, Steinsbekk 2004) were assessed to have low concerns in this domain as the randomisation process and allocation concealment was described in sufficient detail, and no significant baseline differences were noted.

One study (Furuta 2017) was assessed to have some concerns in this domain. This was due to the lack of detail provided about the randomisation process and baseline characteristics. The only baseline characteristic provided was the sex distribution, which was unbalanced across the 2 treatment arms.

Bias due to deviations from intended interventions

All 4 studies (de Lange de Klerk 1993, Furuta 2017, Palm 2017, Steinsbekk 2004) were assessed to have low concerns in this domain. Two studies (De Lange de Klerk 1993) were placebo-controlled, double-blinded studies and one study (Palm 2017) was an open-label study with no blinding.

One study (Steinsbekk 2004) had 4 treatment arms, where only arms 3 (parent-choice) and 4 (placebo) were double-blinded. In this study, the participants and researchers were not blinded to treatment allocation in arms 1 (waitlist) and 2 (homeopath-prescribed). The only deviations noted in this study were non-completion by some participants, however this was not considered related to the trial context. All 4 studies used ITT analysis, and one study (Palm 2017) performed both ITT and per-protocol analysis.

Bias due to missing outcome data

Three studies (de Lange de Klerk 1993, Furuta 2017, Steinsbekk 2004) were assessed to have high concerns in this domain. This was due to the proportion of missing outcome data and the reasons for participant drop-out relating to health status.

One study (Palm 2017) was assessed to have low concerns in this domain. In this study, missing outcome data was addressed using statistical models (Cox model) and sensitivity analysis (Poisson regression).

Bias in measurement of the outcome

Two studies (de Lange de Klerk 1993, Furuta 2017) were assessed to have low concerns in this domain. These studies used validated outcome measures that were consistent between groups, and the outcome assessors were blinded.

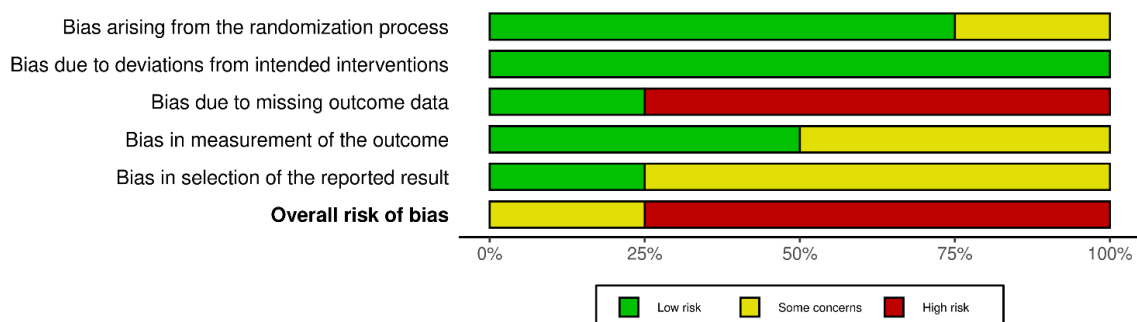
Two studies (Palm 2017, Steinsbekk 2004) were deemed to have some concerns in this domain. Both studies used validated outcome measures that were consistent across treatment groups, however the outcome assessors were not blinded to treatment allocations in Palm 2017, or in treatment arms 1 (waitlist) and 2 (homeopath-prescribed) in Steinsbekk 2004 (study contained 4 treatment arms). It is therefore possible that knowledge of the intervention could have biased outcome measurements, however there was no evidence to suggest this is likely. Note, arms 3 (parent-choice) and 4 (placebo) for Steinsbekk 2004 were double-blinded – therefore this comparison was at low risk of bias for this domain.

Bias in selection of the reported result

Three studies (de Lange de Klerk 1993, Furuta 2017, Palm 2017) were assessed to have some concerns in this domain. In these studies there was no evidence of multiple eligible outcome measurements or analysis, however no pre-specified analysis plans were available for comparison.

One study (Steinsbekk 2004) was assessed to have low concerns in this domain. In this study there was no evidence of multiple eligible outcome measurements or analysis, and the study protocol was available for comparison.

Figure D-4 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Recurrent upper respiratory tract infections



D1.4.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with recurrent upper respiratory tract infections are listed in Table D-2.

Table D-8 Outcomes considered by the NTWC to be critical or important for decision-making: Recurrent upper respiratory tract infections

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	de Lange de Klerk 1993	Furuta 2017	Palm 2017	Steinsbekk 2004
Infection frequency	Number of URTIs, or other validated measure	Critical	Yes	✓ [^]	✓	✓	--
Infection duration	Mean duration of URTI	Critical	Yes	--	--	✓ [^]	✓
Symptom severity	Number of participants experiencing sore throat, or other validated measure	Critical	Yes	✓	--	✓	✓
Quality of life	Wisconsin URTI symptom survey-21 – QoL score	Critical	Yes	✓	--	--	--
Medication use	Antibiotic and/or analgesic use	Critical	Yes	✓	--	✓	✓

Abbreviations: QoL, quality of life; URTI, upper respiratory tract infection

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude, or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

[^] Study data was not able to be extracted

D1.4.3.1 Primary Comparison (vs placebo)

Two RCTs (de Lange de Klerk 1993, Steinsbekk 2004) and one quasi RCT (Furuta 2017) comparing homeopathy with placebo in people with recurrent upper respiratory tract infections (URTIs) were eligible for this comparison. Each study contributed data to at least one outcome domain considered critical for this review.

There were 5 ongoing studies (472 participants) that compared homeopathy with a placebo that could have contributed data to 4 critical outcome domains. Data from one of these studies (232 participants) are available to the study authors but not published. There were also 2 studies awaiting classification (188 participants) that could have contributed data to 2 critical outcome domains.

Infection frequency

Two studies (total 210 participants) reported infection frequency, measured using different methods (de Lange de Klerk 1993, Furuta 2017).

One study (total 170 participants) collected diary recordings of symptoms associated with URTIs experienced over the course of one year (de Lange de Klerk 1993). From the diaries, episodes of infection were identified based on periods when symptoms increased, decreased or plateaued. Results from this study were not able to be included in the meta-analysis as the reported data were incomplete. The study authors reported the mean estimated number of episodes of URTIs was 7.9 in the homeopathy group, compared to 8.4 in the placebo group.

One study (total 40 participants) reported the number of confirmed acute tonsillitis episodes at the end of treatment (4 months) (Furuta 2017). The results suggested homeopathy was associated with fewer episodes of acute tonsillitis when compared with the placebo group (RR 0.40; 95% CI 0.15, 1.07; $p = 0.07$) (GRADE: *Very Low*).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

Infection duration

One study (total 420 participants across all treatment arms) reported infection duration based on symptom diary recordings at the end of treatment (12 weeks) (Steinsbekk 2004). The authors reported median number of days with UTRI symptoms.

The results suggested there is no difference in the median number of days with URTI symptoms; whether the intervention was prescribed by a homeopath (median 8 days [95% CI 4 to 11.6]) or parent-selected (median 9 days [95% CI 16 to 44]) when compared with the placebo group (median 8 days [95% CI 6 to 9]). However, interpretation of median scores is limited as the data are skewed (GRADE: *Very Low*).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

Symptom severity

Two studies (total 437 participants across all treatment arms) reported symptoms severity, measured using a daily symptom score at the end of treatment (range 12 to 52 weeks) (de Lange de Klerk 1993, Steinsbekk 2004).

In one study (total 170 participants) the mean daily symptom score was based on diary recordings of symptoms associated with URIs experienced over the course of one year (de Lange de Klerk 1993). The score had four dimensions relating to symptoms of the nose, ear, and throat, and general symptoms. Respiratory symptoms were given greater weight than general symptoms. The daily score could vary between 0 (no symptoms) to 56 (many symptoms). Results of the study suggest that participants in the homeopathy group had slight worse mean daily symptom scores when compared to the placebo group (MD 0.40; 95% CI -0.02, 0.82; $p = 0.06$). (*GRADE: Very Low*).

The other study (total 267 participants) used a daily symptom diary that recorded 9 symptoms associated with URIs, with a maximum daily total of 11 (higher is worse). Results from this study (Steinsbekk 2004) were reported as a total symptom score, which suggested there was no important difference in symptom severity, whether the intervention was prescribed by a homeopath (median score 24; 95% CI 11.4 to 35.6) or parent-selected (median score 26; 95% CI 16 to 44); as compared to the placebo group (median score 25; 95% CI 14 to 38). However, interpretation of median scores is limited as the data are skewed and cannot be included in the meta-analysis.

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as both studies contributing data were at high risk of bias.

Quality of life

One study (total 170 participants) reported quality of life measured using a modified general well-being questionnaire at the end of treatment (52 weeks) (de Lange de Klerk 1993). The questionnaire covered 4 dimensions of health: sleep, energy, appetite and mood, with answers rated on a 5-point scale (higher is worse). Improvements in quality of life were reported for both the homeopathy and placebo groups, with little to no difference observed (MD 0.64; 95% CI -1.73, 3.01; $p = 0.60$) (*GRADE: Low*).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

Medication use

Two studies (total 503 participants across all treatment arms) reported medication use, based on the number of participants who used antibiotics at the end of treatment (range 12 to 52 weeks) (de Lange de Klerk 1993, Steinsbekk 2004).

Results from one study (de Lange de Klerk 1993) suggested antibiotic use was lower in the homeopathy group over the course of 1-year, but with wide confidence intervals, the result also overlapped with no important difference (RR 0.75; 95% CI 0.53, 1.05; $p = 0.10$).

Results from one study (Steinsbekk 2004) suggested there was no important difference in antibiotic use, whether the intervention was prescribed by a homeopath (RR 0.79; 95% CI 0.38, 1.68; $p = 0.58$) or parent-selected (RR 1.18; 95% CI 0.65, 2.13; $p = 0.74$). To avoid double counting, the homeopath-prescribed results are included in the evidence synthesis.

Taken together (total 377 participants), the pooled results suggested there is little to no difference between the homeopathy and placebo groups (RR 0.76; 95% CI 0.56, 1.03; $p = 0.08$). (*GRADE: Low*).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as both studies contributing data were at high risk of bias.

D1.4.3.2 Secondary Comparison (vs 'inactive' control)

Two RCTs (Palm 2017, Steinsbekk 2004) comparing homeopathy with an inactive control (no intervention) in people with recurrent upper respiratory tract infections were eligible for this comparison. Each study contributed to at least one outcome considered critical for this review.

There was one ongoing study (40 participants, results available but not published) and 2 studies awaiting classification (unknown number of participants, published in a language other than English) that compared homeopathy to an inactive control (no intervention) that could have contributed data to at least one outcome considered critical or important for this review.

Infection frequency

One study (total 256 participants) reported infection frequency, measured by the number of participants with a documented acute throat infection (ATI) within the observation period (8 to 60 weeks) (Palm 2017). Results from Palm 2017 suggested fewer participants in the homeopathy group had an ATI compared with the control group (RR 0.53; 95% CI 0.39, 0.70; $p < 0.0001$). (*GRADE: Moderate*).

The study also estimated the risk of getting an ATI, using a Poisson regression model that accounted for the baseline ATI frequency, as well as the number of days under observation. Here the trialists estimated a rate of 0.59 (95% CI 0.41, 0.86) episodes per year in the homeopathy group, compared with a rate of 1.34 (95% CI 1.08, 1.66) episodes per year in the control group.

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

Infection duration

One study (total 420 participants across all treatment arms) reported infection duration based on symptom diary recordings at the end of treatment (12 weeks) (Steinsbekk 2004). The authors reported median number of days with UTRI symptoms.

The results suggested a difference in the median number of days with URTI symptoms favouring the homeopathy group when compared to the waitlist control group; whether the intervention was prescribed by a homeopath (median difference -5.00; 95% CI -8.70, -1.30; $p = 0.008$) or parent-selected (median difference -4.00; 95% CI -7.48, -0.52; $p = 0.02$). To avoid double counting, the homeopath-prescribed results are included in the evidence synthesis (*GRADE: Very Low*).

The other study (Palm 2017) reported the number of days with occurrence of any tonsillitis specific symptoms, but did not provide any usable data, noting an effect that favoured the homeopathy group ($p < 0.0001$).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

Symptom severity

One study (total 420 participants across all treatment arms) reported symptom severity based on symptom diary recordings at the end of treatment (12 weeks) (Steinsbekk 2004). The daily symptom diary recorded 9 symptoms associated with URTIs, with a maximum daily total of 11 (higher is worse). Results from the study were reported as a total symptom score, but details about the measure are lacking.

The results suggested symptom severity was lower in the homeopathy group when compared with the waitlist control group, whether the intervention was prescribed by a homeopath (MD -20.00; 95% CI -42.98, 2.98; $p = 0.09$) or parent-selected (MD -18.00; 95% CI -42.33, 6.33; $p = 0.15$), but the confidence interval was wide and overlapped with both an important and no important difference. To avoid double counting, the homeopath-prescribed results are included in the evidence synthesis (*GRADE: Very Low*).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

Medication use

Two studies (total 513 participants across all treatment arms) reported medication use, based on the number of participants who used antibiotics during the treatment period (range 12 to 60 weeks) (Palm 2017, Steinsbekk 2004).

One study (Palm 2017) reported the number of participants with a documented ATI requiring antibiotic treatment over the course of the study (60 weeks). The results suggested fewer participants with an ATI in the homeopathy group required antibiotic treatment (26 out of 50) compared with those in the control group (59 out of 87) (RR 0.77; 95% CI 0.57, 1.04; $p = 0.09$). When the outcome was measured based on the total episodes of ATI, an effect favouring homeopathy was observed (34/92 vs 110/89; RR 0.63; 95% CI 0.47, 0.85; $p = 0.002$).

Results from one study (Steinsbekk 2004) suggested there was little to no difference in antibiotic use when compared to the waitlist control group; whether the intervention was prescribed by a homeopath (RR 0.58; 95% CI 0.28, 1.20; $p = 0.13$) or parent-selected (RR 0.77; 95% CI 0.43, 1.40; $p = 0.39$). To avoid double counting, the homeopath-prescribed results are included in the evidence synthesis.

Taken together (total 306 participants), the pooled results suggest fewer participants in the homeopathy group require antibiotics when compared with the control (no intervention, waitlist) groups (RR 0.74; 95% CI 0.56, 0.97; $p = 0.03$) (*GRADE: Low*).

In a sensitivity analysis that examined the impact of one RCT (Steinsbekk 2004) judged to be at a high risk of bias, the overall direction of effect did not substantially change, but increased doubts about the clinical importance of the effect (see results for Palm 2017).

Quality of life

One study (total 256 participants) reported quality of life, measured using a 5-point rating scale during the study (60 weeks) (Palm 2017). Details of the scale were lacking, with the authors simply noting that participants in the homeopathy group rated their quality of life higher than participants in the control group ($p < 0.001$). No other data were provided.

D1.4.3.3 Tertiary Comparison (vs other)

No studies comparing homeopathy with 'other' interventions in people with recurrent upper respiratory tract infections were identified.

D1.5 Recurrent infections (genitourinary)

D1.5.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-9. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-9 Overview of PICO criteria of included studies: Recurrent urinary tract infections

STUDY ID	STUDY DESIGN	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
No studies found						
Homeopathy versus inactive control*						
Pannek 2019 (44)	Quasi RCT	Recurrent UTIs	Individualised homeopathy, oral + consultation with a homeopath	Inactive control (no intervention)	Standard prophylaxis	Infection frequency; HRQoL
Homeopathy versus 'other' intervention**						
Witt 2009 (45)	RCT	Recurrent vulvovaginal candidiasis	Individualised homeopathy, oral + consultations with a homeopath	Active control (oral itraconazole) OR Active control (oral itraconazole PLUS vaginal lactobacilli tablet)	None reported	No priority outcome domains reported

Abbreviations: HRQoL, health-related quality of life; RCT, randomised controlled trial; UTI, urinary tract infection.

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D1.5.2 Risk of bias per item

The risk of bias for each item in the included studies for recurrent infections (genitourinary) is described below and shown graphically in Figure D-5 (details are provided in Appendix E).

Bias arising from the randomisation process

One study (Pannek 2019) was assessed to have high risk of bias in this domain due to abandonment of the randomisation process prior to the end of the enrolment period. This was a consequence of the open study design, where participants declined informed consent due to knowledge of treatment allocation. Some slight baseline imbalances were noted.

One study (Witt 2009) was assessed as having some concerns due to the lack of information provided on allocation sequence concealment and no baseline characteristics shown.

Bias due to deviations from intended interventions

One study (Pannek 2019) was assessed to have low concerns in this domain. The only deviations reported were non-completion by some participants, but this was deemed not due to the trial context.

One study (Witt 2009) was assessed as having some concerns in this domain due to the lack of information provided on blinding, and the use of a per-protocol analysis method.

Bias due to missing outcome data

One study (Pannek 2019) was assessed as having some concerns in this domain due to missing outcome data with no adjustments presented. Reasons for participant drop-out were not provided, therefore it is possible that knowledge of treatment allocation contributed to drop-out.

One study (Witt 2009) was assessed to have high concerns in this domain due to the very high proportion (53%) of missing data. Health status was among the reasons for participant drop-out, which was generally evenly distributed across treatment groups.

Bias in measurement of the outcome

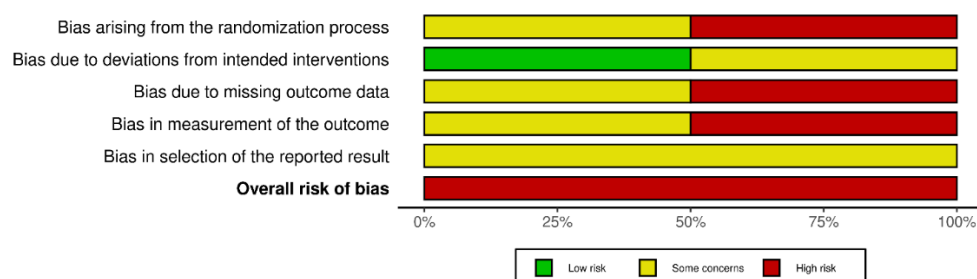
One study (Pannek 2019) was assessed as having some concerns in this domain. This was an open trial where it was possible that knowledge of the intervention could bias the assessment of the outcomes.

One study (Witt 2009) was assessed to have high concerns in this domain. This study did not provide any information on blinding. It was therefore not able to be determined if the outcome assessors were aware of the interventions received, and whether this may have influenced the assessment of the outcome.

Bias in selection of the reported result

Both studies (Pannek 2019 and Witt 2009) were assessed to have some concerns in this domain due to the lack of pre-specified analysis plans.

Figure D-5 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – recurrent urinary tract infections



D1.5.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with recurrent infections (genitourinary) are listed in Table D-10.

Table D-10 Outcomes considered by the NTWC to be critical or important for decision-making: Recurrent urinary tract infections

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Pannek 2019	Witt 2009
Infection frequency	Number of infections (symptomatic or total)	Critical	Yes	✓	--
Health-related quality of life	EQ-5D	Critical	No	✓^	--

Abbreviations: EQ-5D, EuroQol five dimensions questionnaire.

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results
- ? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study.
- ^ Study data not in extractable form

D1.5.3.1 Primary Comparison (vs placebo)

No studies were identified that compared homeopathy to placebo in people with recurrent genitourinary infections. There were no ongoing studies and no studies awaiting classification.

D1.5.3.2 Secondary Comparison (vs 'inactive' control)

One quasi RCT (Pannek 2019) comparing homeopathy to an inactive control (no intervention) in people with recurrent urinary tract infections (UTIs) was eligible for this comparison and contributed data to 2 outcomes considered critical for this review.

There were no ongoing studies and no studies awaiting classification that compared homeopathy with an inactive control (no intervention, waitlist, usual care).

Infection frequency

One quasi RCT (35 participants) reported infection frequency, measured both prospectively and retrospectively (Pannek 2019). The prospective method involved a self-reporting questionnaire and confirmatory UTI dipstick test. The number of infections experienced throughout the 12-month study period was recorded. The retrospective method involved collecting the participants' recollections about the number of UTIs experienced at the end of the study.

Results from the study suggested that there was no difference in the number of UTIs experienced in the homeopath group when compared with the control group (RR 0.27; 95% CI 0.05, 1.36; $p = 0.11$) (GRADE: Very Low).

No sensitivity analysis was conducted that examined the impact of RCTs judged to be at a high risk of bias, as only one study contributed data.

Health-related quality of life

One quasi RCT (Pannek 2019) (35 participants) measured health-related quality of life using the EQ-5D at the end of treatment (12 months). The EQ-5D is a standardised health-related quality of life measure which measures an individual's health state on 5 scales: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The study authors reported that homeopathy did not have a significant effect on quality of life ($p > 0.9$). The results were not able to be included in the meta-analysis as the study did not report any usable data (such as means, SD, 95% CI).

D1.5.3.3 Tertiary Comparison (vs other)

One RCT (Witt 2009) (48 participants) compared homeopathy with 'other' interventions in people with recurrent infections (genitourinary). This study did not contribute to any priority outcome domains considered critical for this review.

D2 Mental, behavioural or neurodevelopmental disorders

D2.1 Anxiety

D2.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-11. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-11 Overview of PICO criteria of included studies: Anxiety

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Baker 2003 (46)	RCT	Performance test anxiety (students, ≥ 50 on the Benson RTA)	Non-individualised homeopathy, oral (Argentum nitricum) traditionally OR radionically prepared ^	Placebo	None reported	Anxiety severity
Bonne 2003 (47)	RCT	Generalised Anxiety Disorder	Individualised homeopathy, oral	Placebo	None reported	Anxiety severity Depression severity Psychological symptoms Psychological wellbeing
Dimpfel 2016 (48)	Quasi RCT	Performance test anxiety (adults, > 60 on the PAF-S)	Non-individualised homeopathy, oral	Placebo	None reported	Cognitive function Emotional function
Foy-Nux 2018 (49)	Quasi RCT	Acute anxiety (predental, children)	Non-individualised homeopathy, oral (combination)	Placebo	Sedation (inhaled nitrous oxide)	Anxiety severity Behaviour
Parewa 2021 (50)	RCT	Generalised Anxiety Disorder	Individualised homeopathy, oral	Placebo	Monthly psychological counselling	Anxiety severity
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus other 'active' interventions**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial; RTA: Revised Test Anxiety; PAF-S: Prüfungsangstfragebogen Sie (Test Anxiety questionnaire)

^ Study included 3 groups. Both the traditionally and radionically prepared are included in the evidence synthesis.

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D2.1.2 Risk of bias per item

The risk of bias for each item in the included studies for anxiety is described below and shown graphically in Figure D-6 (details are provided in Appendix E).

Bias arising from the randomisation process

One study (Parewa 2021) was assessed to be at low risk of bias. The other 4 studies (Baker 2003, Bonne 2003, Dimpfel 2016, Foy-Nux 2018) were assessed to have some concerns due to bias arising from the randomisation process. Concerns were primarily related to lack of information regarding the allocation concealment process or the method of generating the random sequence.

Bias due to deviations from intended interventions

Two studies (Bonne 2003, Parewa 2021) were judged to be at low risk for this domain as intent-to-treat or modified intent-to-treat analysis was used. Two studies (Dimpfel 2016, Baker 2003) were judged to have some concerns. For one study (Dimpfel 2016) this was due to the lack of information presented regarding the method of analysis used and for the second study (Baker 2003) this was due to using a per-protocol analysis. One study, (Fux-Noy 2018), was judged to be at high risk for this domain due to the large dropout as a result of non-compliance (almost 50%) and using an as-treated analysis approach.

Bias due to missing outcome data

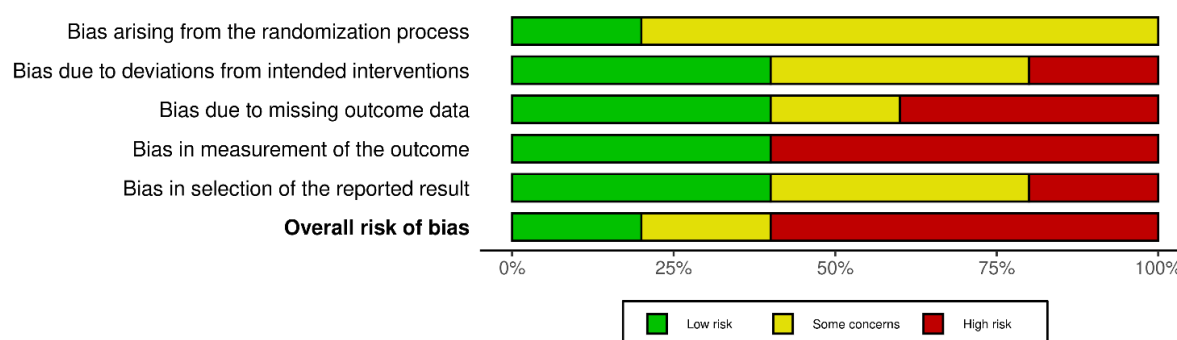
Two studies (Dimpfel 2016, Parewa 2021) were assessed to be at low risk for this domain. In one study (Dimpfel 2016) outcome analysis included data on all participants of the study. In the second study, (Parewa 2021), 10% of participants dropped out of the study and missing values were estimated using regression. The study authors have not presented evidence to show results were not biased by this missing outcome data, however equal numbers of participants dropped out in both treatment groups. One study (Bonne 2003) was judged to have some concerns regarding missing outcome data as data on more than 5% of participants was missing. Last observation carried forward (LOCF) was completed and these results did not differ significantly from the base case, however LOCF should not be assumed to correct for missingness. Two studies (Baker 2003 and Fox-Nuy 2018) were deemed to have a high risk of bias for this domain as in one (Baker 2003) more than 5% of participant data was missing and this was not accounted for and in the other, (Foy-Nux 2018), data from more than 50% of participants was missing, mainly due to non-compliance, and this was not accounted for.

Bias in measurement of the outcome

Two studies (Bonne 2003 and Parewa 2021) were judged to be at low risk of bias for this domain as the outcome assessors remained blinded and the measures used were validated for this population. Three studies (Baker 2003, Fox-Nuy 2018, Dimpfel 2016) were deemed to be high risk for this domain as outcome measurement could have differed between treatment groups (Baker 2003, Dimpfel 2016) and the outcome measures used are not validated in the assigned population (Fox-Nuy 2018 and Dimpfel 2016). Information on the blinding of outcome assessors was not provided by study authors for the 3 studies (Baker 2003, Fox-Nuy 2018, Dimpfel 2016).

Bias in selection of the reported result

Two studies (Bonne 2003, Parewa 2021) were deemed to be low risk for this domain as a pre-specified data analysis plan was provided. Two studies (Baker 2003, Fox-Nuy 2018) were assessed to have some concerns regarding bias in this domain as no pre-specified analysis plan was available. The final study, (Dimpfel 2016), was deemed to be at high risk for this domain as no pre-specified analysis plan was available and interpretation of the results were dependent on recording conditions.

Figure D-6 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – anxiety

D2.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with anxiety are listed in Table D-12.

Table D-12 Outcomes considered by the NTWC to be critical or important for decision-making: Anxiety

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Baker 2003	Bonne 2003	Dimpfel 2016	Foy-Nux 2018	Parewa 2021
Anxiety symptoms	HAM-A (or other validated measure)	Critical	Yes	✓	✓	--	--	✓
Depression	HAM-D	Critical	Yes	--	✓	--	--	--
Emotional function	BSI (or other validated measure)	Critical	Yes	--	✓	--	--	--
Physical function	PROMIS physical function	Critical	No	--	--	--	--	--
Quality of life	WHO-QoL-BREF	Critical	No	--	--	--	--	--

Abbreviations: BSI: Brief Symptom Inventory; HAM-A, Hamilton Anxiety Rating Scale; HAM-D: Hamilton Depression Rating Scale; RTA: Revised Test Anxiety Scale

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

D2.1.3.1 Primary Comparison (vs placebo)

Three RCTs (Baker 2003, Fux-Noy 2018, Parewa 2021) and 2 quasi RCTs (Bonne 2003, Dimpfel 2016) comparing homeopathy with placebo in people with anxiety were eligible for this comparison. The studies contributed data to 3 outcomes considered critical or important for this review.

There were 2 studies awaiting classification (total 98+ participants³) and 2 ongoing studies (total 92 participants) comparing homeopathy with placebo that could have contributed data to the anxiety severity outcome (see Appendix C6 for details).

³ Numbers not reported in one study.

Anxiety severity

Three studies (150 participants) reported anxiety symptoms measured with the Hamilton Anxiety Rating Scale (HAM-A) or the Revised Test Anxiety scale (RTA) at the end of treatment (range 4 days to 3 months) (Baker 2003, Bonne 2003, Parewa 2021).

The HAM-A is widely used in both clinical and research settings to measure the severity of anxiety symptoms. The scale consists of 14 items each scored on a scale from 0 (not present) to 4 (severe) to yield a total score from 0 to 56 where a higher score indicates more severe anxiety. Remission, which represents complete or near complete symptom resolution, including resolution of functional impairment, is defined as a HAM-A total score of 7 or less (51). The results from 2 studies (106 participants) (Bonne 2003, Parewa 2021) showed no difference in anxiety symptoms in the homeopathy group compared to the placebo group (MD -2.25; 95% CI -5.92, 1.43; $p = 0.23$; $I^2 = 32\%$).

The RTA is a 20-item scale which measures 4 dimensions: worry, tension, test-irrelevant thinking and bodily symptoms. Items are graded on a 4-point Likert scale, with total score ranging from 20 to 80 and a higher score indicating worse test anxiety (52). No MCID for the RTA was identified.

Results from one study (Baker 2003) suggested there was no difference between the homeopathy and placebo groups, whether the preparation was traditionally prepared (44 participants) (MD 5.42; 95% CI -2.05, 12.88; $p = 0.16$) or radionically prepared (MD 0.55; 95% CI -4.36, 5.47; $p = 0.83$). Results from the traditional homeopathy arm used in the meta-analysis as this aligns with how the intervention would be delivered in Australian practice.

Pooled results suggest that there is no difference between treatment groups comparing homeopathy with placebo (SMD -0.05; 95% CI -0.66, 0.56; $p = 0.87$; $I^2 = 71\%$) (*GRADE: Low*).

In a sensitivity analysis examining the impact of one study (Baker 2003) at high risk of bias, the estimate of effect did not materially change, with no difference in anxiety symptoms in the homeopathy group compared to the placebo group (SMD -0.28; 95% CI -0.94, 0.38; $p = 0.41$; $I^2 = 65\%$).

Depression severity

One study (44 participants) reported depression measured with the Hamilton depression rating scale (HAM-D) at the end of treatment (10 weeks) (Bonne 2003). The HAM-D measures the severity of depressive symptoms experienced in the previous week and consists of 17 or 21-items scored on a 3- or 5-point scale⁴. Individual scores are summed (range 0 to 52) with a higher score indicating a greater level of depressive symptoms. Reported MCIDs range between 3 and 8 points, with a score above 20 indicative of clinical depression and scores below 7 considered normal (53, 54).

The results from one study suggest there was little to no difference between treatment groups at the end of treatment comparing homeopathy with placebo (MD 1.50; 95% CI -2.16, 5.16; $p = 0.42$) (*GRADE: Low*).

No sensitivity analysis was performed examining the impact of studies at high risk of bias as only one study contributed data to this outcome.

⁴ It is assumed the original 17-item version is used in the included study.

Emotional function

One study (44 participants) reported emotional function measured with the Brief Symptom Inventory (BSI) at the end of treatment (10 weeks) (Bonne 2003). The BSI is a 53-item self-report measure of psychological distress developed for use in clinical settings. It covers nine dimension relating to emotional function: somatization, obsession-compulsion, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism; and three global indices of distress: global severity index, positive symptom distress index, and positive symptom total (55). Items are scored on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely) in relation to the previous 7 days, with total scores interpreted by comparison to age-appropriate norms (55).

The results from one study (Bonne 2003) suggested there was no difference between homeopathy and placebo groups at the end of treatment (MD 0.00; CI -0.08, 0.08; $p = 1.00$). (*GRADE: Low*).

No sensitivity analysis was performed examining the impact of studies at high risk of bias as only one study contributed data to this outcome.

The study also reported emotional function measured with the Psychological General Wellbeing Index at end of treatment. Study authors report no difference between the homeopathy and placebo group for this measure (see Appendix F2).

D2.1.3.2 Secondary Comparison (vs inactive control)

There were no RCTs comparing homeopathy with inactive control (no intervention, waitlist, usual care) in people with anxiety that were eligible for this comparison.

D2.1.3.3 Tertiary Comparison (vs other)

There were no RCTs comparing homeopathy with 'other' interventions in people with anxiety that were eligible for this comparison.

There is one ongoing study that is eligible for this comparison (comparing homeopathy with pharmacotherapy [benzodiazepine]). The study is listed as complete, with results not yet available.

D2.2 Depression

D2.2.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-13. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-13 Overview of PICO criteria of included studies: Depression

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Adler 2011 (56, 57)	RCT	Adults with severe major depression	Individual homeopathy, oral	Placebo	homeopathic case history taking OR conventional case history taking [^]	Depression symptoms HRQoL (physical and mental component scores)
Katz 2005 (58)	RCT	Adults experiencing a moderate major depressive episode	Individualised homeopathy, oral (selected from list of 30 commonly prescribed)	Placebo	None reported	Depression symptoms, HRQoL
Homeopathy versus inactive control*						
Viksvveen 2014 (59-61)	RCT	Adults with major depressive disorder	Individualised homeopathy (offered)	No intervention	Usual care	Depression symptoms
Homeopathy versus other 'active' interventions**						
Adler 2009 (62-64)	RCT	Adults with moderate to severe major depression (single or recurrent episode)	Individual homeopathy, oral	Pharmacotherapy (fluoxetine)	All participants took a dummy of their respective comparator	Depression symptoms
Katz 2005 (58)	RCT	Adults experiencing a moderate major depressive episode	Individualised homeopathy, oral (selected from list of 30 commonly prescribed)	Pharmacotherapy (fluoxetine)	None reported	Depression symptoms HRQoL

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: HRQoL, health-related quality of life; RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

[^]The conventional and homeopathic case histories differed in the time used for the semi-standardised questionnaire and the onsite patient-doctor interaction. The homeopathic case history was a more extensive conversation (60–90 minutes), in which the patient was asked to speak about different aspects in the questionnaire, including stressful life events, development and details of psychological and physical symptoms, and information which was discussed as needed. This was different in the shorter, more conventional case history which took around 30 minutes. The same questionnaire was read in silence by the attending physician, who asked questions only to elucidate information that was unclear, resembling a more conventional case taking, conducted by a general physician, when the patient already has a psychiatric diagnosis of depression. The follow-up differed in the time used to assess the remaining symptoms, 10 or 30 minutes, for the conventional or the homeopathic case history, respectively. In the latter case, interpersonal or ongoing stressful life events were also extensively assessed.

D2.2.2 Risk of bias per item

The risk of bias for each item in the included studies for depression is described below and shown graphically in Figure D-7 (details are provided in Appendix E).

Bias arising from the randomisation process

Three studies (Adler 2009, Katz 2005, Viksveen 2014) were at low risk of bias in this domain. One study (Adler 2011) had some concerns of bias relating to baseline differences between groups and lack of allocation concealment.

Bias due to deviations from intended interventions

All studies (Adler 2009, Adler 2011, Katz 2005, Viksveen 2014) were assessed to have low risk of bias for this domain.

Bias due to missing outcome data

All studies (Adler 2009, Adler 2011, Katz 2005, Viksveen 2014) were at high risk of bias in this domain relating to large proportions of missing data considered likely to be due to the true value of the outcome (Adler 2009), with no reasons provided for drop out (Adler 2011, Katz 2005) and which could be related to the true outcome value (Viksveen 2014). The proportion of missing data from the end analyses of studies ranged from 16% (Adler 2011) to 90% (Katz 2005).

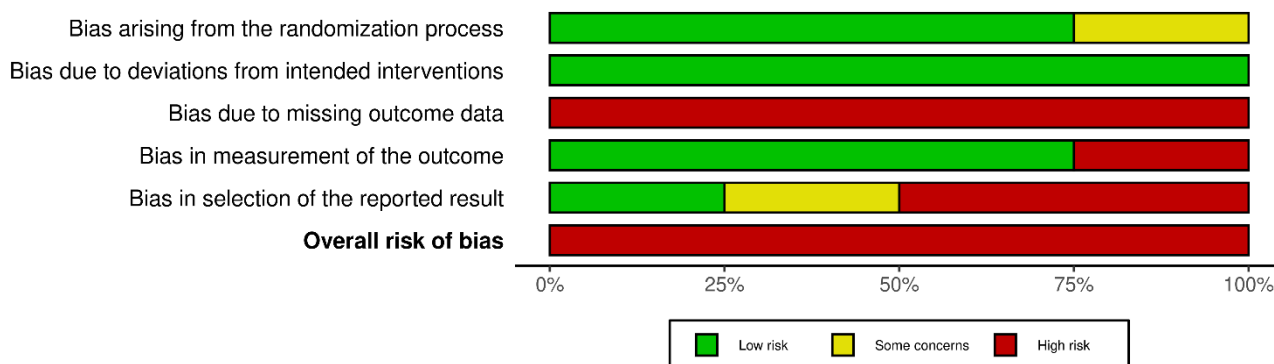
Bias in measurement of the outcome

Three studies (Adler 2009, Adler 2011, Katz 2005) were at low risk of bias for this domain. One study (Viksveen 2014) was assessed to have high risk of bias for this domain, arising from the lack of blinding of participants and the recruitment strategy. Due to the study design, those who elected to uptake the offer of homeopathy were considered likely to differentially report their outcomes due to a belief in the efficacy of treatment.

Bias in selection of the reported result

One study (Adler 2011) was judged to be low risk of bias and one study (Adler 2009) was judged to have some concerns for this domain due to the lack of information provided on pre-specified analysis. Two studies (Katz 2005, Viksveen 2014) were assessed to have high risk of bias for this domain arising from the changes in outcomes measured. Outcomes excluded could have contributed data to the analysis.

Figure D-7 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – depression



D2.2.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with depression are listed in Table D-14.

Table D-14 Outcomes considered by the NTWC to be critical or important for decision-making: Depression

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Adler 2009	Adler 2011	Katz 2005	Viksvveen 2014
Depression symptoms	BDI, HAM-D, MADRS, PHQ-9, CGI	Critical	Yes	✓	✓	X	✓
Psychological distress	Kessler K10 (or similar)	Critical	No	--	--	--	--
Emotional functioning	SF-12 mental component score (or similar)	Critical	Yes	--	✓	?	--
Physical functioning	SF-12 physical component score (or similar)	Critical	Yes	--	✓	?	--
Health-related quality of life	SF-12 (or similar)	Critical	No	--	--	?	? ^a

Abbreviations: BDI, Beck depression inventory; CGI, clinical global impression; HAM-D: Hamilton depression rating scale; MADRS, Montgomery-Asberg depression rating scale; PHQ-9, patient health questionnaire-9; SF-12: 12-item short form

a. Quality of life reported in the study protocol, but authors later state the outcome was removed to reduce patient burden.

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

D2.2.3.1 Primary Comparison (vs placebo)

Two studies (Adler 2011, Katz 2005) comparing homeopathy with placebo in people with depression were eligible for this comparison. One study (Adler 2011) contributed data to 3 outcomes considered critical or important for this review. The second study (Katz 2005) was a pilot study involving a total of 6 participants and did not report any usable data.

There were no studies awaiting classification and 5 ongoing studies (total 518 participants) comparing homeopathy to placebo that were eligible for this comparison. One ongoing study (total 40 participants) was considered likely to be complete, but results are not reported (see Appendix C6 for details).

No sensitivity analyses were performed examining the impact of studies at high risk of bias as only one study contributed data.

Depression symptoms

Two studies (total 48 participants) reported depression symptoms measured using the Hamilton depression rating scale (HAM-D) at end of treatment (range 6 to 12 weeks) (Adler 2011, Katz 2005).

The HAM-D measures the severity of depressive symptoms experienced in the previous week and consists of 17 or 21-items scored on a 3- or 5-point scale⁵. Individual scores are summed (range 0 to 52) with a higher score indicating a greater level of depressive symptoms. Reported MCIDs range between 3 and 8 points, with a score above 20 indicative of moderate to severe clinical depression and scores below 7 considered normal (53, 54).

⁵ The 17-item scale was used by Adler 2011

Results from one study (total 44 participant) suggested that there was no important difference in depressive symptom at the end of treatment comparing homeopathy with placebo (MD 2.33; 95% CI –0.51, 5.17; $p = 0.11$; $I^2 = 0\%$) (*GRADE: Very Low*). The observed results were similar, whether the intervention was delivered with homeopathic case history taking (MD 3.10; 95% CI –0.84, 7.04) or conventional case history taking (MD 1.50; 95% CI –2.60, 5.60).

Adler 2011 also measured depression symptoms using the Beck Depression Inventory (BDI). The result of this analysis using the BDI outcome did not differ substantially from the HAM-D (see Appendix F2).

Results from one study (Katz 2005) were not able to be included in the meta-analysis as the study authors did not report SD. The study authors did not observe any difference in depression symptoms between the homeopathy and placebo groups.

Emotional functioning

One study (total 44 participants) reported emotional functioning measured using the SF-12 health survey at end of treatment (6 weeks) (Adler 2011). One other study measured this outcome but did not report any data (Katz 2005).

The SF-12 measures the impact of one's health on everyday life across 8 domains. The mental component summary (MCS) score is derived by aggregating individual scores of 4 domains associated with emotional wellbeing: vitality, social functioning, role emotional, and mental health; and is summarised on a scale from 0 (worse) to 100 (best). A higher score indicates improved emotional wellbeing. The MCID for the SF-12 in people with depression is unknown but is estimated to be around 2 to 4 points for the general population (i.e. ~ 0.5 of the SD) (65).

The results suggested there was no important difference in emotional functioning between the homeopathy and placebo groups at the end of treatment (MD –1.90; 95% CI –9.14, 5.35; $p = 0.61$; $I^2 = 0\%$) (*GRADE: Very Low*). The observed results were better when the intervention was delivered with homeopathic case history taking (MD –4.30; 95% CI –13.82, 5.22) than conventional case history taking (MD 1.40; 95% CI –9.76, 12.56), but the difference is not clinically important.

Physical functioning

One study (total 44 participants) reported physical functioning measured using the SF-12 health survey at end of treatment (6 weeks) (Adler 2011). One other study measured this outcome but did not report any data (Katz 2005).

The SF-12 physical component summary (PCS) score is derived by aggregating individual scores of 4 domains associated with physical wellbeing: general health, physical functioning, role-physical and bodily pain; and is summarised on a scale from 0 (worse) to 100 (best). A higher score indicates improved physical functioning. The MCID for the SF-12 in people with depression is unknown but is estimated to be around 2 to 4 points for the general population (i.e. ~ 0.5 of the SD) (65).

The results suggested there was no important difference in physical functioning between the homeopathy and placebo groups at the end of treatment (MD –4.77, 95% CI, –11.29, 1.75 $p = 0.15$; $I^2 = 14\%$) (*GRADE: Very Low*). The observed results were better when the intervention was delivered with homeopathic case history taking (MD –7.30; 95% CI –14.65, 0.05) than conventional case history taking (MD –0.40; 95% CI –10.53, 9.73), but with low participant numbers the difference is not important.

D2.2.3.2 Secondary Comparison (vs 'inactive' control)

One study (Viksveen 2014) comparing homeopathy with inactive control (no intervention) in people with depression was eligible for this comparison and contributed data for one outcome considered critical or important for this review.

Depression symptoms

One study (total 566 participants) reported depression symptoms measured using the 9-item patient health questionnaire (PHQ-9) at the end of treatment (52 weeks) (Viksveen 2014). The PHQ-9 is used to assess patients for the presence and intensity of depressive symptoms over the past 2 weeks and is scored on a 4 points scale from 0 (not at all) to 3 (nearly every day). Total score ranges from 0 to 27 (higher is worse), with scores below 4 considered minimal to no depression and those 20 or above associated with severe depression (66). A reduction of 5-points is reported to be clinically important (67).

The study authors did not report individual group post-treatment or change from baseline scores but noted an effect favouring participants offered homeopathy when compared with the control group (MD -1.4 95% CI -2.5, -0.2, $p = 0.015$) (*GRADE: Moderate*), but the result was not clinically important. A similar result was observed when only participants who accepted the offer for homeopathic care were included in the analysis (MD -2.4 95% CI -4.0, -0.9, $p = 0.002$)

No sensitivity analysis was performed examining the impact of studies at high risk of bias as only one study contributed data.

D2.2.3.3 Tertiary Comparison (vs active intervention)

Two studies (Adler 2009, Katz 2005) (total 99 participants) comparing homeopathy with active intervention (anti-depressant) in people with depression were eligible for this comparison. Data from these studies are presented in Appendix F2 Supplementary outcome data.

D2.3 Neurodevelopmental disorders

D2.3.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-15. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-15 Overview of PICO criteria of included studies: Neurodevelopmental disorders

STUDY ID	STUDY DESIGN	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Dhawale 2014 (68)	Quasi RCT	Children with dyslexia and dysgraphia	Individualised homeopathy	Placebo	Consultations with a homeopath and remedial education	Study did not report any priority outcomes
Frei 2005 (69-73)	RCT	Children with ADHD	Individualised homeopathy	Placebo	Consultations with a homeopath	ADHD symptoms
Jacobs 2005 (74)	RCT	Children with ADHD	Individualised homeopathy	Placebo	Consultations with a homeopath	ADHD symptoms
Lamont 1997 (75)	Quasi RCT	Children with ADHD	Individualised homeopathy	Placebo	Consultations with a homeopath	Study did not report any priority outcomes
Oberai 2013 (76)	RCT	Children with ADHD	Individualised homeopathy	Placebo	Consultations with a homeopath	ADHD symptoms Quality of life
Strauss 2000 (77)	Quasi RCT	Children with ADHD	Non-individualised (Selenium-Homaccord)	Control (not specified) ^	Standard medical care (+/- methylphenidate)	ADHD symptoms
Homeopathy versus inactive control*						
Fibert 2015 (78-83)	RCT	Children with ADHD	Individualised homeopathy, attended up to 8 consultations with a homeopath	Inactive control (no intervention)	None reported	ADHD symptoms
Homeopathy versus 'other' intervention**						
Fibert 2015 (78-83)	RCT	Children with ADHD	Individualised homeopathy, attended up to 8 consultations with a homeopath	Diet, participants attended up to 8 consultations with a nutritional therapist	None reported	ADHD symptoms

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: ADHD, attention deficit hyperactivity disorder; RCT, randomised controlled trial.

^ Comparator group described both as placebo and control throughout the report but not clear if placebo or no intervention.

Participants were initially stratified into those receiving methylphenidate or not, then randomised to the homeopathy/control arms

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D2.3.2 Risk of bias per item

The risk of bias for each item in the included studies for neurodevelopmental disorders is described below and shown graphically in Figure D-8 (details are provided in Appendix E).

Bias arising from the randomisation process

Three studies (Fibert 2015, Frei 2005, Jacobs 2005) were assessed to have low concerns in this domain as the randomisation process and allocation concealment was described in sufficient detail, and no significant baseline differences were noted.

Three studies (Dhawale 2014, Oberai 2013, Strauss 200) were assessed as having some concerns in this domain. In Dhawale 2014 and Strauss 2000, details on the randomisation process, allocation concealment and baseline characteristics were not provided in sufficient detail. In Oberai 2013, the randomisation process was sufficiently described, and baseline characteristics were comparable, however, no information was provided on the allocation concealment.

One study (Lamont 1997) was assessed as having high concerns in this domain. This was due to the randomisation process being described only as alternate assignment of groups, and no information provided on allocation concealment or baseline characteristics.

Bias due to deviations from intended interventions

Three studies (Frei 2005, Jacobs 2005 and Oberai 2013) were assessed to have low concerns in this domain. In these studies, the participants and the people delivering the interventions were blinded to treatment allocations. ITT analysis was performed in Frei 2005 and Jacobs 2005, and a modified ITT analysis was performed in Oberai 2013.

Four studies (Dhawale 2014, Fibert 2015, Lamont 1997 and Strauss 2000) were assessed as having some concerns in this domain. In Fibert 2015, the study was not blinded, so participants and the people delivering the interventions were aware of treatment allocations. Potential deviations from the intended intervention included the option for participant withdrawal, however there is nothing to suggest this affected the outcome. This study used both ITT and per protocol analyses.

In Dhawale 2014, Lamont 1997 and Strauss 2000, the participants were blinded to treatment allocations. The people delivering the interventions were also blinded in all studies except Dhawale 2014. The method of analysis was not reported in these studies and was not able to be determined from the results presented.

Bias due to missing outcome data

One study (Jacobs 2005) was assessed to have some concerns in this domain. In this study there was some evidence of analysis performed to correct for bias, however the method of analysis was not described. The study had missing outcome data due to participant drop-out, however the reasons for drop-out were not provided.

Six studies (Dhawale 2014, Fibert 2015, Frei 2005, Lamont 1997, Oberai 2013, Strauss 2000) were assessed to have high concerns in this domain. Of these, 2 studies (Dhawale 2014 and Strauss 2000) did not report if there was missing outcome data, so it could not be determined if the outcome was affected. Two studies (Lamont 1997 and Oberai 2013) reported there was some missing outcome data however, no details were provided as to the reasons for participant non-completion. Two studies (Fibert 2015 and Frei 2005) reported missing outcome data and provided reasons for participant drop-out, which included reasons that were related to health status.

Bias in measurement of the outcome

Three studies (Frei 2005, Jacobs 2005, Strauss 2000) were assessed to be at low risk of bias for this domain. In these studies the outcome assessors were blinded, and valid outcome measures were used.

Two studies (Fibert 2015, Oberai 2013) were assessed to have some concerns in this domain. In these studies, the outcome assessors were aware of treatment allocations. It was therefore possible that knowledge of the intervention could have influenced the measurement of the outcome, however, there was no evidence to suggest this was likely.

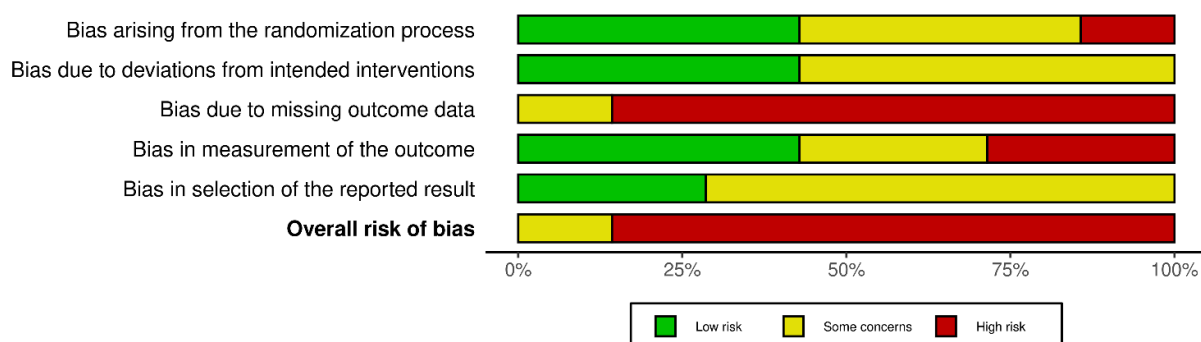
Two studies (Dhawale 2014, Lamont 1997) were assessed to have high concerns in this domain. In Lamont 1997, the study authors noted it was possible that the outcome assessor could have influenced the outcome of the trial in favour of the hypothesis. In Dhawale 2014, no details were provided on how the outcomes were measured. It was therefore not able to be determined if the method of measuring the outcome was appropriate and consistent across the treatment groups.

Bias in selection of the reported result

Two studies (Fibert 2015, Frei 2005) were assessed as having low concerns in this domain, as pre-specified analysis plans were indicated in the studies, and there was no evidence of bias in the selection of the reported results.

Five studies (Dhawale 2014, Jacobs 2005, Lamont 1997, Oberai 2013, Strauss 2000) were assessed to have some concerns in this domain due to the lack of pre-specified analysis plans. There was no evidence of bias in the selection of the reported results in these studies.

Figure D-8 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Neurodevelopmental disorders



D2.3.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with attention deficit hyperactivity disorder are listed in Table D-16.

Table D-16 Outcomes considered by the NTWC to be critical or important for decision-making: Neurodevelopmental disorders

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison ?	Dhawale 2014	Fibert 2015	Frei 2005	Jacobs 2005	Lamont 1997	Oberai 2013	Strauss 2000
ADHD symptoms	Conner's global index or Conner's rating scale	Critical	Yes	--	✓	✓^	✓	--	✓	✓^
Behaviour	Child Behaviour Checklist (or other validated measure)	Critical	No	--	--	--	--	--	--	--
Emotional function	Profile of mood states	Critical	No	--	--	--	--	--	--	--
HRQoL	ADHD quality of life scale (or similar)	Critical	No	--	✓	--	--	--	--	--

Abbreviations: ADHD, attention deficit hyperactivity disorder; HRQoL, health-related quality of life

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

^ Study data was not able to be extracted

D2.3.3.1 Primary Comparison (vs placebo)

Three RCTs (Frei 2005, Jacobs 2005, Oberai 2013) and 3 quasi RCTs (Dhawale 2014, Lamont 1997, Strauss 2000) comparing homeopathy with placebo in people with neurodevelopmental disorders were eligible for comparison. The studies contributed data to one outcome domain considered critical or important for this review.

There were no studies awaiting classification and 3 ongoing studies (total 323 participants) comparing homeopathy with placebo in children with ADHD or autism spectrum disorder. One study (112 participants) was considered likely to be completed (but results not reported) and could have contributed data to at least one outcome (see Appendix C6).

ADHD symptoms

Four studies (Frei 2005, Jacobs 2005, Oberai 2013, Strauss 2000) reported ADHD symptoms measured using a version of the Conner's rating scale at the end of treatment (range 6 weeks to 12 months).

Two studies (total 104 participants) measured ADHD symptoms using the Conner's Global Index (CGI) (parent and teacher rated). The CGI is a 10-item scale focused on ADHD symptoms that are considered the most important: temper, outbursts, excitable, impulsive, overactive, cries often, inattentive, fidgeting, disturbs other children, easily frustrated, fails to finish things, moods change quickly. The scale ranges from 0 (never) to 3 (very often). Total scores are standardised against population norms on a scale from 0 (best) to 100 (worse), with a T-score higher than 60 indicating symptoms of ADHD.

In one study (Frei 2005), a reduction of 5 points from the pre-treatment value was considered clinically relevant, however results were not able to be included in the meta-analysis as the study did not report means, SD, or 95% CI. The trialists reported an effect favouring the homeopathy treatment group (parent-rated; MD -1.67, 95% CI NR; $p=0.0479$). Data for teacher-rated results were not reported.

Results from Jacobs 2005 (total 43 participants) suggested there was no difference in ADHD symptoms between the homeopathy group and the placebo groups at the end of treatment (18 weeks); whether they were parent- (MD 1.77; 95% CI –6.34, 9.88; $p = 0.67$) or teacher-rated (MD 4.72; 95% CI –2.11, 11.55; $p = 0.18$) (*GRADE: Very Low*).

Two studies (total 104 participants) used the Conner's Parents Rating Scale – Revised: Short (CPRS-R:S), which was designed for repeated and/or brief assessment of symptoms relevant to ADHD and related disorders (84). The 27-item parent short form (CPRS-R:S) comprises 4 subscales/indices (oppositional, cognition, hyperactivity, and ADHD index), with respondents asked to rate behaviour that has been problematic over the preceding month using a 4-point Likert scale. Higher scores indicate worse symptoms, with males typically receiving higher scores than females. An MCID for the CPRS-R:S has not been established.

Pooled results from 2 studies (Jacobs 2005, Oberai 2013) suggested the ADHD symptom scores for each of the 4 subscales/indices were lower in the homeopathy treatment group when compared with the placebo group at the end of treatment, but there was substantial statistical heterogeneity and the confidence intervals overlapped with no benefit (*GRADE: Very Low*).

- CPRS-R:S Oppositional: MD –8.01; 95% CI –25.73, 9.72; $p = 0.38$; $I^2 = 93\%$
- CPRS-R:S Cognition: MD –5.69; 95% CI –26.24, 14.87; $p = 0.59$; $I^2 = 96\%$
- CPRS-R:S Hyperactivity: MD –10.02; 95% CI –35.15, 15.11; $p = 0.43$; $I^2 = 96\%$
- CPRS-R:S ADHD index: MD –7.54; 95% CI –25.76, 10.68; $p = 0.42$; $I^2 = 95\%$

One study (total 20 participants) was not able to be included in the data synthesis for this outcome as it did not report any usable data (Strauss 2000), but authors noted an effect favouring homeopathy.

In a sensitivity analysis examining the impact of one study (Oberai 2013) at high risk of bias, the results for each of the 4 subscales/indices shifted to favouring the placebo group:

- CPRS-R:S Oppositional: MD 1.40; 95% CI –6.86, 9.66; $p = 0.74$
- CPRS-R:S Cognition: MD 5.08; 95% CI –2.45, 12.61; $p = 0.19$
- CPRS-R:S Hyperactivity: MD 3.05; 95% CI –5.46, 11.56; $p = 0.48$
- CPRS-R:S ADHD index: MD 2.00; 95% CI –4.92, 8.92; $p = 0.57$

D2.3.3.2 Secondary Comparison (vs 'inactive' control)

One RCT (Fibert 2015) (total 83 participants) comparing homeopathy with an inactive control (no intervention) in people with ADHD was eligible for comparison. The studies contributed to 4 outcome domains considered critical for this review.

There was one ongoing study (151 participants) that compared homeopathy with an inactive control (no intervention) that could have contributed data to one critical outcome domain (results published after the search date (85)). There were no studies awaiting classification.

No sensitivity analyses were performed examining the impact of studies at high risk of bias as only one study contributed data.

ADHD symptoms

One study (Fibert 2015) reported ADHD symptoms measured using the Conner's global index (CGI) (parent- and teacher-rated) at the end of treatment (12 months). The CGI is a 10-item scale focused on ADHD symptoms that are considered the most important: temper, outbursts, excitable, impulsive, overactive, cries often, inattentive, fidgeting, disturbs other children, easily frustrated, fails to finish things, moods change quickly. Scored on a 4-point Likert scale, the CGI is reported as a total score or 2 sub-scores measuring restlessness/impulsivity (7 items) and emotional lability (3 items). Higher scores indicate worse ADHD symptoms. Total scores are standardised against population norms on a scale from 0 (best) to 100 (worse), with a T-score higher than 60 indicating symptoms of ADHD.

Results from Fibert 2015 suggested there was little to no difference between the homeopathy group and the inactive control group at the end of treatment (12 months) when measured using the parent-rated form (*GRADE: Very Low*):

- CGI – total: MD 2.03; 95% CI –0.72, 4.78; $p = 0.15$
- CGI – restless-impulsive: MD 1.47; 95% CI –0.56, 3.50; $p = 0.16$
- CGI – emotional lability: MD 0.55; 95% CI –0.55, 1.65; $p = 0.33$

Teacher-rated scores were not included here as too few reports were returned for meaningful analysis.

Quality of life

One study (Fibert 2015) reported health-related quality of life (HRQoL) measured using the Child Health Utility 9D Index (CHU-9) at the end of treatment (12 months).

The CHU-9 is a generic preference-based measure of HRQoL suitable for children aged 7 to 17 years. It consists of 9 short questions completed by the child that relate to the following: worried, sad, pain, tired, annoyed, schoolwork, sleep, daily routine, and activities. There are 5 response levels ranging from 0 (It is never a problem) to 4 (It is almost always a problem) (86). Total scores are standardised on a scale from 0 (dead) to 1 (perfect health).

Results from Fibert 2015 suggested there was little to no difference between the homeopathy group and the inactive control group at the end of treatment (12 months) (MD –0.01; 95% CI –0.07, 0.05; $p = 0.76$). (*GRADE: Very Low*).

D2.3.3.3 Tertiary Comparison (vs other)

Two RCTs comparing homeopathy with 'other' interventions in people with ADHD were eligible for this comparison and contributed data for one critical outcome domain. Data from these studies are presented in **Appendix F2** Supplementary outcome data.

D3 Sleep-wake disorders

D3.1 Insomnia and sleep problems

D3.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-17. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-17 Overview of PICO criteria of included studies: Insomnia and sleep problems

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Harrison 2013 (87)	Quasi RCT	Insomnia (chronic, primary)	Non-individualised homeopathic combination, oral	Placebo	None reported	Pre-sleep arousal, Sleep latency
James 2019 (88)	RCT	Insomnia (chronic)	Individualised homeopathy, oral	Placebo	Sleep hygiene encouraged	Insomnia severity, Insomnia symptoms
Naude 2010 (89)	Quasi RCT	Insomnia (chronic, primary)	Individualised homeopathy, oral	Placebo	None	Insomnia severity, Insomnia symptoms
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus 'other' intervention**						
Jong 2016 (90-92)	RCT	Sleep disorders (children)	Non-individualised homeopathic combination, oral	Glycine, oral	None reported	Sleep symptoms severity

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D3.1.2 Risk of bias per item

The risk of bias for each item in the included studies for insomnia and sleep disorders is described below and shown graphically in Figure D-9 (details are provided in Appendix E).

Bias arising from the randomisation process

Two studies (James 2019, Jong 2016) were assessed at low risk of bias for this domain as they provided information regarding the generation of the randomisation sequence, method of allocation concealment, and there were no baseline imbalances between the treatment arms suggestive of an issue with randomisation. Two studies (Harrison 2013, Naude 2010) were assessed to have some concerns as the method of randomisation was judged to be not truly random and allocation concealment was not adequately reported.

Bias due to deviations from intended interventions

Two studies (James 2019, Jong 2016) were judged at low risk of bias for this domain. While Jong 2016 did not involve blinding of participants or research staff, there were no deviations from the intended intervention

that were considered likely due to the trial context, and both studies employed an appropriate method of analysis (ITT). One study (Naude 2010) was judged to have some concerns in this domain due to the potentially inappropriate method of analysis which involved the exclusion of an otherwise eligible participant for non-compliance with the intended intervention. This was not considered likely to have a substantial impact on the results as only one participant (3% of the total study) was excluded for this reason. One study (Harrison 2013) was judged at high risk of bias for an inappropriate method of analysis which excluded participants due to non-compliance. This was determined to potentially have a significant impact on the results due to the proportion of participants who were excluded (2 participants [12%] in the placebo arm and at least one [total unknown] in the homeopathy arm).

Bias due to missing outcome data

Two studies (James 2019, Jong 2016) were assessed at low risk of bias for this domain as they either had outcome data available for nearly all participants, or missing outcome data was appropriately adjusted for in the analysis. One study (Naude 2010) was judged to have some concerns due to the amount of missing data, however it was not considered likely that missingness was due to the true value of the outcome. One study (Harrison 2013) was judged to be at high risk of bias in this domain due to the amount of missing outcome data which was not adjusted for, and which was considered likely to be due to the true value of the outcome.

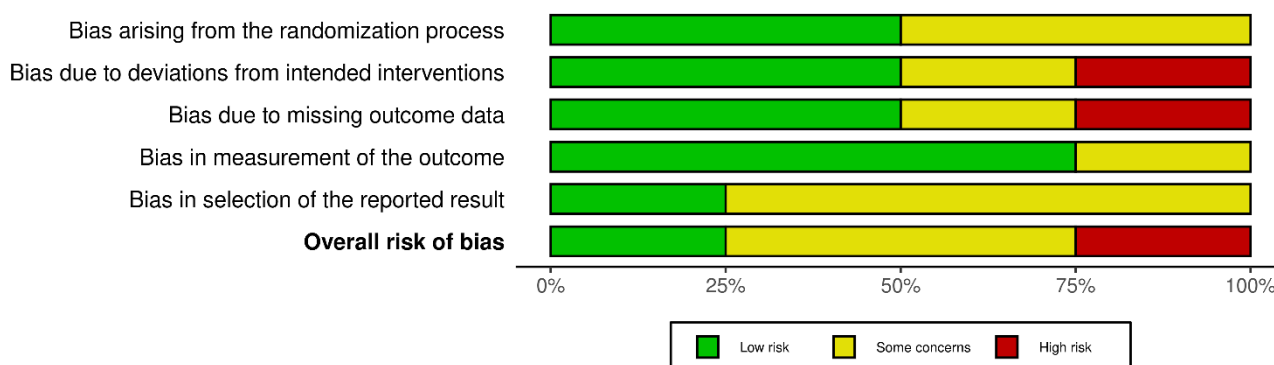
Bias in measurement of the outcome

Three studies (Harrison 2013, James 2019, Naude 2010) were judged at low risk of bias for this domain as the outcomes were measured appropriately, and outcome assessors were blinded to intervention status. One study (Jong 2016) was judged to have some concerns for this domain, as self-reported outcomes were assessed by non-blinded participants. It was not considered likely that reporting of the outcome would be influenced by knowledge of the intervention.

Bias in selection of the reported result

One study (James 2019) was judged at low risk of bias for this domain. While no pre-specified analysis plan was included, it was noted that a published protocol was available as part of a postgraduate thesis and all reported outcomes correspond with the clinical trial registry. Three studies (Harrison 2013, Jong 2016, Naude 2010) were judged to have some concerns for this domain due to lack of available information on pre-specified analysis plans.

Figure D-9 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Insomnia and sleep problems



D3.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with insomnia are listed in Figure D-18.

Table D-18 Outcomes considered by the NTWC to be critical or important for decision-making: Insomnia and sleep problems

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Harrison 2013	James 2019	Naude 2010
Insomnia severity	Insomnia Severity Index (or other validated measure)	Critical	Yes	--	✓	✓ [^]
Sleep quality / sleep duration	Pittsburgh Sleep Quality Index (or other validated measure)	Critical	Yes	--	✓	✓
Sleep onset latency	Polysomnography (or other validated measure)	Critical	Yes	✓	✓	?
Quality of life	SF-36 (or other validated measure)	Critical	No	--	--	--

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

[^] Study data was not able to be interpreted

D3.1.3.1 Primary Comparison (vs placebo)

One RCT (James 2019) and 2 quasi RCTs (Harrison 2013, Naude 2010) comparing homeopathy with placebo in people with insomnia were eligible for this comparison. The studies contributed data to 3 outcomes considered critical or important for this review.

There was one study awaiting classification (44 participants) that compared homeopathy with placebo in people with insomnia that could have contributed data to one outcome, but the study was published in a language other than English (see Appendix C4). There was one ongoing study (60 participants) open for recruitment that will contribute data to 2 outcomes when complete.

No sensitivity analyses were performed examining the impact of studies at high risk of bias as only one study contributed data to each outcome.

Insomnia severity

Two studies (90 participants) reported insomnia severity measured with the Insomnia Severity Index (ISI) or the Sleep Impairment Index (SII) at the end of treatment (range: 4 weeks to 3 months) (James 2019, Naude 2010).

The ISI is a 7-item questionnaire assessing the nature, severity and impact of insomnia, with the focus being on subjective feelings about insomnia symptoms. Each question is summed to give a total score that ranges from 0 to 28, where a higher score is indicative of worse insomnia. Scores are categorised as follows: 0-7, no clinical insomnia; 8-14, subclinical insomnia; 15-21, clinical insomnia (moderate); 22-28, clinical insomnia (severe). A cut-off score of 10 has been found to maximise sensitivity and specificity in a community sample (93). In a clinical sample of people seeking treatment for insomnia, an improvement of 8.4 points corresponded to a moderate improvement in insomnia (93).

Results from one study (60 participants) (James 2019) showed a difference in ISI score between the homeopathy group and the control group at the end of treatment, however this difference is not considered clinically meaningful (MD -2.70; 95% CI -4.73, -0.67; $p = 0.009$). (GRADE: Low)

The SII is a 7-item scale that rates a person's perceived severity of sleep onset, sleep maintenance, and early morning awakening problems in the past 2 weeks; as well as current interference with daytime functioning; noticeability of impairment caused by the sleep problem; distress/concern caused by the sleep problem; and satisfaction with current sleep pattern (94, 95). Total scores range from 5 to 35 (higher is worse). The MCID for the SII is unknown.

The results from one study (total 30 participants) (Naude 2010) were not able to be interpreted and therefore were not included in the data synthesis. The authors did not report SD for the SII, and the reported values do not correlate with expected values. The study authors report a significant benefit in favour of homeopathy but it is not clear if the effect is clinically meaningful.

Sleep duration

Two studies (90 participants) reported sleep duration measured using a sleep diary at the end of treatment (range: 4 weeks to 3 months) (James 2019, Naude 2010).

A sleep diary is a daily record of an individual's sleep pattern, including hours of sleep, sleep latency, sleep interruptions and sleep quality. One study (James 2019) reports sleep duration as total hours of sleep per night and one study (Naude 2010) reports sleep duration as total hours of sleep per week (more is better). No MCID was identified in the literature, however a Cochrane review assessing the effectiveness of acupuncture for insomnia has previously noted that a difference of less than one hour per night might not be clinically relevant (96).

Results from one study (60 participants) (James 2019) showed no significant difference in hours of sleep per night between the homeopathy and placebo groups (MD -0.10; 95% CI -0.76, 0.56; $p = 0.77$). (GRADE: Low)

Naude 2010 (30 participants) did not report SD and could not be included in the meta-analysis. Study authors report a significant difference in sleep hours per week between the homeopathy group and placebo group (6 hours per week as reported by study authors). As this converts to a difference of less than one hour per night, it may not be clinically meaningful.

Sleep onset latency

Two studies (88 participants) reported sleep onset latency measured using a sleep diary at the end of treatment (range: 4 weeks to 3 months) (James 2019, Harrison 2013).

A sleep diary is a daily record of an individual's sleep pattern, including hours of sleep, sleep latency, sleep interruptions and sleep quality. One study (James 2019) reports sleep onset latency as the number of minutes taken to fall asleep (more is worse). One study (Harrison 2013) reports sleep latency as a categorical variable based on the time to fall asleep: from 0 (0-15 mins) to 4 (60+ mins), with a higher score indicating worse sleep latency. No MCID was identified in the literature, however a Cochrane review assessing the effectiveness of acupuncture for insomnia has previously noted that a difference of less than 30 minutes in sleep onset latency might not be clinically relevant (96).

Results from one study (60 participants) (James 2019) showed no difference in sleep onset latency between the homeopathy and placebo groups (MD -22.20; 95% CI -45.18, 0.78; $p = 0.06$). (GRADE: Low)

One study (Harrison 2013) (total 28 participants) does not report SD, as such it was not included in the meta-analysis. Study authors report a significant difference in sleep latency between the homeopathy group and placebo group, however it is unclear whether this difference is clinically meaningful.

D3.1.3.2 Secondary Comparison (vs 'inactive' control)

There were no studies identified which compared homeopathy with inactive control.

D3.1.3.3 Tertiary Comparison (vs other)

One RCT comparing homeopathy with 'other' interventions in children with sleep impairment was eligible for this comparison and contributed data for 2 outcomes. Data from this study are presented in Appendix F2 Supplementary outcome data.

D4 Diseases of the nervous system

D4.1 Headache disorders

D4.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-19. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-19 Overview of PICO criteria of included studies: Headache disorders

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Straumsheim 1997 (97, 98)	RCT	Migraine	Individualised homeopathy, oral	Placebo	None reported	Headache severity, Headache frequency, Medication use
Gaus 1992(99-102)	RCT	Headache	Individualised homeopathy, oral	Placebo	None reported	Headache severity, Headache frequency, Medication use
Whitmarsh 1997 (103)	Quasi RCT	Migraine	Individualised homeopathy, oral	Placebo	None reported	Headache severity, Headache frequency,
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus 'other' intervention**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D4.1.2 Risk of bias per item

The risk of bias for each item in the included studies for people with headache disorders is described below and shown graphically in Figure D-10 (details are provided in Appendix E).

Bias arising from the randomisation process

One study (Straumsheim 1997) was assessed at low risk of bias for this domain as they provided information regarding the generation of the randomisation sequence, method of allocation concealment, and there were some differences in baseline characteristics, which were not considered likely due to issues with randomisation.

One study (Gaus 1992) was assessed to have some concerns for this domain, as baseline differences for previous treatment of headaches were different between the homeopathy and placebo groups. One study (Whitmarsh 1997) was assessed as high risk of bias, as no information was provided regarding the generation of the randomisation sequence, and the method of allocation concealment. Additionally the mean migraine attack frequency was significantly higher in the placebo group.

Bias due to deviations from intended interventions

Two studies (Gaus 1992, Whitmarsh 1997) were assessed at low risk of bias for this domain, as the participants and research staff were both blinded to the intervention, and appropriate analysis was undertaken. One study (Straumsheim 1997) was assessed to have some concerns for this domain. The pharmacist responsible for distributing the intervention had access to the randomisation code. Additionally, analysis of the results excluded 5 participants; one patient experienced no migraine in the month prior to treatment, 2 became pregnant, one became hypertensive, and one was lost to follow-up, it was not specified which treatment group these participants withdrew from.

Bias due to missing outcome data

One study (Gaus 1992) was assessed at low risk of bias for this domain, as missing outcome data was appropriately adjusted for in the analysis. Two studies (Straumsheim 1997, Whitmarsh 1997) were assessed at high risk of bias for this domain. Missing data was not adjusted for in either study, and it was considered likely the missing data could be due to the true value of the outcome.

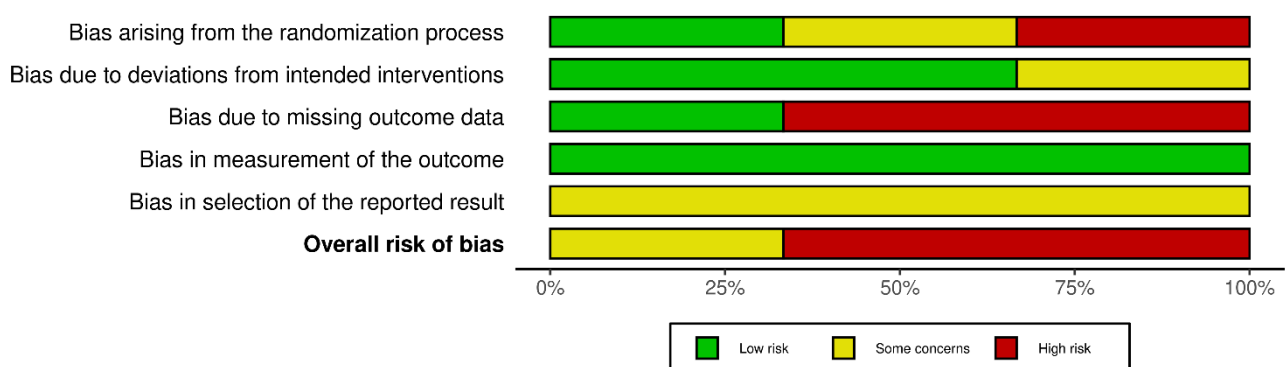
Bias in measurement of the outcome

Three studies (Straumsheim 1997, Gaus 1992, Whitmarsh 1997) were judged at low risk of bias for this domain as the outcomes were measured appropriately. In all studies participants who were blinded to the study intervention self-reported outcomes

Bias in selection of the reported result

All studies were assessed to have some concerns for this domain, as there was no pre-specified analysis plan available. There was no evidence of inappropriate analysis or reporting of results in any of the included studies.

Figure D-10 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Headache disorders



D4.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with headache and migraine are listed in Figure D-20.

Table D-20 Outcomes considered by the NTWC to be critical or important for decision-making: Headache disorders

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Straumsheim 2000	Gaus 1992	Whitmarsh 1997
Headache frequency	% days with headache, Number of days per month	Critical	Yes	✓ [^]	✓ [^]	✓ [^]
Headache severity	VAS or Rating Scale	Critical	Yes	--	--	✓
Pain	VAS	Critical	Yes	✓	✓	--
Migraine duration	hours	Critical	Yes	--	✓	--
Headache impact	HIT-6 MIDAS	Critical	No	--	--	--
Medication use	Self-reported diary	Critical	Yes	✓	✓	--
Quality of life	SF-36	Critical	No	--	--	--

Abbreviations: HIT-6, Headache Impact Test 6; MIDAS, Migraine Disability Assessment; SF-36, short-form 36; VAS, visual analogue scale.

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

[^] Study data was not able to be extracted (or included in the synthesis)

D4.1.3.1 Primary Comparison (vs placebo)

Two RCTs (Gaus 1992, Straumsheim 1997) and one quasi RCT (Whitmarsh 1997) comparing homeopathy to placebo in people with headache or migraine were eligible for this comparison and contributed data for 5 outcomes considered critical or important for this review.

There was one study awaiting classification (total 50 participants) and one ongoing study (12 participants) that compared homeopathy with placebo and could have contributed data to at least one outcome considered critical or important by the NTWC (see Appendix C6).

Headache/migraine frequency

Three studies (226 participants) reported headache/migraine frequency per month based on diary records and assessed the change from baseline at the end of treatment (range 3 to 4 months).

Data from all three studies was incomplete and not able to be included in the meta-analysis. Data was synthesised qualitatively for analysis.

One study (Gaus 1992) reported a median change in headache frequency of one per month in both the homeopathy and placebo groups, but statistical analysis was not conducted by the trialists. One study suggested (Straumsheim 1997) that migraine frequency decreased in both the homeopathy and placebo groups, with the change from baseline being higher in the placebo group, but the difference was not significant ($p = 0.54$). The final study (Whitmarsh 1997) reported a 19.02% reduction in migraine frequency in the homeopathy group compared with a 16.46% reduction in the placebo group ($p = 0.83$).

Overall, two out of three studies suggested there was no significant difference between homeopathy and placebo. The third study reported a tendency favouring placebo but did not conduct statistical analysis (*GRADE: Very Low*).

Headache/migraine severity

One study (total 63 participants) reported headache/migraine severity measured based on participant-rated categories (mild, moderate, severe); with results reported as the change in frequency during the trial of headaches of each category of severity by treatment group (from baseline to end of treatment [4 months]) (Whitmarsh 1997)

The study authors reported relative decreases in each category of severity measured, but in the absence of any baseline or end of treatment numbers, the data were not able to be included in the analysis (see Appendix F2 Supplementary outcome data). (*GRADE: Very Low*).

Pain intensity

Two studies (total 163 participants) reported headache/migraine pain intensity measured using a 100 mm visual analogue scale (VAS) at the end of treatment (range 3 to 4 months) (Gaus 1997, Straumsheim 1997). A VAS is a unidimensional measure of pain where participants are asked to rate their pain on a scale from 0 (best) to 100 mm (worst). An MCID in people with headache disorders has not been established.

Data from both studies were incomplete and not able to be included in the meta-analysis.

Results from one study (98 participants) (Gaus 1992) suggested there was no difference in the median change in VAS scores between treatment groups, but statistical analysis was not conducted by the trialists. Data from one study (Straumsheim 1997) was not able to be included in the meta-analysis, as reported data were incomplete. The authors noted a reduction in pain intensity in both groups (54% versus 42%) comparing homeopathy with placebo ($p = 0.08$). Overall, data from two studies suggest no difference in pain intensity between homeopathy and placebo (*GRADE: Very Low*).

Headache duration

One study (98 participants) reported headache duration (hours) based on diary records and assessed the change from baseline at the end of treatment (3 months) (Gaus 1992).

Data was incomplete and not able to be included in the meta-analysis.

Results suggested there was no difference in the median headache duration between treatment groups (*GRADE: Very Low*).

Medication use

Two studies (total 163 participants) reported medication used at the end of treatment (range 3 to 4 months), but data were incomplete and not able to be use in the data synthesis.

One study (98 participants) reported a reduction in the mean daily dose (mg) across 8 drugs, in both treatment groups (Gaus 1992). The difference between groups was not significant ($p = 0.16$).

One study (65 participants) reported a 52% reduction in medication use in the homeopathy group and a 42% reduction in the placebo group, noting the result was not significant ($p = \text{not reported}$) (Straumsheim 1997).

Overall, two studies reported no difference in medication use between homeopathy and placebo (*GRADE: Very Low*).

D4.1.3.2 Secondary Comparison (vs 'inactive' control)

There were no RCTs found and no ongoing studies that compared homeopathy to inactive control (no intervention, waitlist, usual care) in people with headache disorders.

There were 2 studies awaiting classification (127+ participants), which could have contributed data to 4 outcomes considered critical by the NTCWC (See Appendix C6).

D4.1.3.3 Tertiary Comparison (vs other)

There were no RCTs found comparing homeopathy with 'other' interventions in people with headache disorders.

D5 Diseases of the respiratory system

D5.1 Asthma

D5.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-21. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#).

Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-21 Overview of PICO criteria of included studies: Asthma

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Lewith 2002 (104)	Quasi RCT	Asthma (with allergy to dust mite)	Non-individualised homeopathy, oral	Placebo	Usual care (concurrent medications remained unchanged)	Asthma symptoms Pulmonary function Medication use
Qutubudd in 2019 (105)	RCT	Asthma, bronchial	Individualised homeopathy, oral	Placebo	Usual care Both groups received consultation with a homeopath	Asthma symptoms Pulmonary function
Reilly 1994 (106, 107)	RCT	Asthma (allergic)	Individualised homeopathy, oral	Placebo	Usual care Both groups received consultation with a homeopath	Asthma symptoms Pulmonary function
White 2003 (108)	RCT	Asthma, bronchial (age 5 to 15 yrs)	Individualised homeopathy	Placebo	Usual care Both groups received consultation with a homeopath	Pulmonary function Quality of life Medication use
Homeopathy versus inactive control						
Thompson 2008 (109, 110)	Quasi RCT	Asthma, bronchial (age 7 to 14 yrs)	Individualised homeopathy, oral, attended one initial consultation with a homeopath	Inactive control (no intervention)	Usual care	Asthma symptoms Pulmonary function

Topcu 2010 (111-113)	RCT	Asthma, bronchial	Individualised homeopathy, attended 6-12 consultations with a homeopath	Inactive control (no intervention) OR Reflexology [^]	Usual care	Asthma symptoms Pulmonary function Quality of life
Homeopathy versus 'other' intervention**						
No studies found [^]						

Abbreviations: RCT, randomised controlled trial.

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

[^] One study (Topcu 2010) included 2 comparator groups. The inactive control is considered in Secondary Comparison. Data for the reflexology group is presented in [Appendix F2](#).

D5.1.2 Risk of bias per item

The risk of bias for each item in the included studies for asthma is described below and shown graphically in Figure D-11 (details are provided in Appendix E).

Bias arising from the randomisation process

All 6 studies (Lewith 2002, Qutubuddin 2019, Topcu 2010, White 2003, Thompson 2008 and Reilly 1994) were assessed to have low risk of bias for this domain. Five studies (Lewith 2002, Qutubuddin 2019, Topcu 2010, White 2003 and Reilly 1994) reported in sufficient detail their randomisation processes, allocation concealment and baseline characteristics. One study (Thompson 2008) did not provide details of their randomisation process, however details of allocation concealment and baseline characteristics were documented and considered sufficient.

Bias due to deviations from intended interventions

All 6 studies (Lewith 2002, Qutubuddin 2019, Topcu 2010, White 2003, Thompson 2008 and Reilly 1994) were assessed to have low risk of bias for this domain. Four studies (Lewith 2002, Qutubuddin 2019, White 2003 and Reilly 1994) were double-blind, placebo-controlled trials that did not report any deviations from the intended interventions due to the trial context. ITT analysis was specified and conducted in these studies.

In 2 studies (Topcu 2010 and Thompson 2008), participants and those delivering the intervention were not blinded to treatment allocation. Non-completion by some participants was noted, but this was not considered to be due to the trial context. ITT analysis was specified and conducted in these studies.

Bias due to missing outcome data

Two studies (Topcu 2010 and Thompson 2008) were assessed to have some concerns in this domain due to the proportion of missing outcome data, which was not considered likely to be related to the true value of the outcome. Four studies (Qutubuddin 2019, Reilly 1994, Lewith 2002 and White 2003) were judged at high concerns in this domain. This was due to the proportion of missing outcome data which was considered likely to be related to the true value of the outcome. In each of these studies, authors report that drop out in some participants was related to worsening health status or lack of improvement.

Bias in measurement of the outcome

Four studies (Lewith 2002, Qutubuddin 2019, White 2003 and Reilly 1994) were assessed to have low risk of bias for this domain. These studies used valid outcome measures that did not differ between treatment groups, and the outcome assessors were blinded to treatment allocation.

Two studies (Topcu 2010 and Thompson 2008) were assessed as having some concerns in this domain. In these studies, participants were aware of their treatment allocation, so self-reported outcome measures may have been influenced by this knowledge.

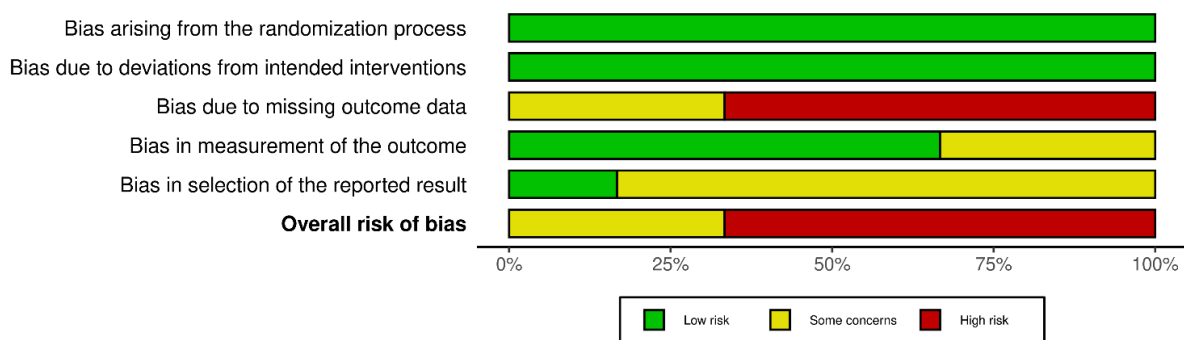
Bias in selection of the reported result

Five studies (Lewith 2003, Topcu 2019, White 2003, Thompson 2008 and Reilly 1994) were assessed to have some concerns in this domain as they did not have pre-specified analysis plans available. In 2 studies (Reilly 1994, White 2003), study authors did not present data on all outcomes measured. However, the authors report outcomes in-text. Data was not able to be extracted for analysis, however this was not considered to raise risk of bias to high overall.

In 2 studies (White 2003 and Reilly 1994), there was evidence of selective reporting of outcomes. In these studies, some self-reported outcomes were measured but the results were not shown. It was not able to be determined if these were trial deviations, as there were no study protocols available for comparison.

One study (Qutubuddin 2019) was assessed to have low risk of bias for this domain. Data produced in this study were analysed in accordance with a published, pre-specified analysis plan.

Figure D-11 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – asthma



D5.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with asthma are listed in Table D-22.

Table D-22 Outcomes considered by the NTWC to be critical or important for decision-making: Asthma

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Lewith 2002	Qutubuddin 2019	Reilly 1994	Thompson 2008	Topcu 2010	White 2003
Asthma symptoms	ACQ or other validated measure	Critical	Yes	✓ [^]	✓	✓	✓	✓	--
Pulmonary function	FEV ₁ /FVC ratio, FEV ₁ , PEFR or other validated measure	Critical	Yes	✓	✓	✓	✓	✓ [^]	✓
Quality of life	AQLQ or CAQ	Critical	Yes	✓	--	--	✓	✓	✓
Hospitalisation	Exacerbation requiring hospitalisation	Critical	No	--	--	--	✓	--	--
Medication use	Use of inhalers	Critical	No	✓ [^]	--	--	✓	✓ [^]	✓

Abbreviations: ACQ, asthma control questionnaire; AQLQ, asthma quality of life questionnaire; CAQ, childhood asthma questionnaire; FEV₁, forced expiratory rate in first second; PEFR, peak expiratory flow rate, FVC, forced vital capacity

✓ A study result is available for inclusion in the synthesis

- X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators
- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results
- ? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study
- ^ Study data was not able to be extracted

D5.1.3.1 Primary Comparison (vs placebo)

Three RCTs (Qutubuddin 2019, Reilly 1994, White 2003) and one quasi RCT (Lewith 2002) comparing homeopathy with placebo in people with asthma were eligible for comparison (total 455 participants). The studies contributed data to 4 of 5 outcomes considered critical or important for this review.

There were 4 studies awaiting classification (170+ participants, noting the number of participants was not clear in 2 studies) that compared homeopathy with placebo that could have contributed data to at least 2 critical outcome domains (pulmonary function and medication use) (see Appendix C6).

One ongoing study (total 102 participants) that compared homeopathy with placebo was found that could contribute to one outcome domain critical for this review, but the study is not yet recruiting (see Appendix C4).

No sensitivity analysis was performed examining the impact of studies at high risk of bias as all studies contributing data were at high risk of bias.

Asthma symptoms

Three studies (Lewith 2002, Qutubuddin 2019, Reilly 1994) reported asthma symptoms measured using the Asthma Control Questionnaire (ACQ) or a Visual Analogue Scale (VAS) at end of treatment (range: 4 weeks to 6 months).

One study (Qutubuddin 2019) (140 participants) reported asthma symptoms measured using the ACQ at end of treatment (6 months). The ACQ consists of 7 questions relating to asthma symptoms, activity limitation, use of medications and pulmonary function, measured on a 7-point scale (0 = totally controlled, 6 = extremely poorly controlled). A change in score of 0.5 or more is considered the MCID for this outcome measure (114). The results suggested a clinically meaningful improvement in ACQ scores favouring the homeopathy group (MD -0.70; 95% CI -0.78, -0.62; $p < 0.0001$).

Two studies (Reilly 1994, Lewith 2002) (226 participants) reported asthma symptoms using a VAS. The VAS is a validated tool used for the subjective measurement of symptom severity on a continuous scale (usually from 0 to 100), with higher scores indicating greater severity. Results from one study (Reilly 1994) (total 28 participants) suggested a reduction in VAS scores (change from baseline) favouring the homeopathy group at the end of treatment (4 weeks) (MD -15.00; 95% CI -23.60, -6.40; $p = 0.0006$). Results from one study (Lewith 2002) (total 202 participants) were not able to be included in the meta-analysis as study authors did not report usable data (means, SD etc.), but noted there was no important difference between the groups at the end of treatment (16 weeks).

Pooled results (total 168 participants) suggest there is a reduction in asthma symptoms in the homeopathy group compared with placebo (SMD -2.03; 95% CI -3.48, -0.59; $p = 0.006$; $I^2 = 89\%$) (GRADE: Very Low). There is substantial statistical heterogeneity, some of which may be explained by differences in the populations (bronchial or allergic asthma) and treatments received (individualised homeopathy).

Pulmonary function

Four RCTs (total 449 participants) reported pulmonary function measured using spirometry tests at the end of treatment (range 4 weeks to 12 months) (Lewith 2002, Qutubuddin 2019, Reilly 1994, White 2003). A spirometry test is a type of lung function test that measure how much air you can forcefully exhale and is used to confirm a diagnosis of asthma or assess asthma severity (115).

One study (Qutubuddin 2019) reported the ratio between the forced expiratory volume in 1 second (FEV₁) and forced vital capacity (FVC). Normal values for the FEV₁/FVC ratio typically range between 75% to 85%, depending on sex and age, with a 5-point percentage lower than normal suggestive of airflow obstruction (lower is worse) (I16). An MCID for a change in FEV₁/FVC has not been established, but the normalisation of the ratio, secondary to an improvement in FEV₁ could be considered clinically important (I17).

Results from this study (140 participants) suggested pulmonary function was higher in the homeopathy group at the end of treatment (6 months) (87.1%) when compared with the placebo group (82.5%) (MD -0.05; 95% CI -0.08, -0.01; $p = 0.007$), noting the FEV₁/FVC ratio declined in both groups and remained within the normal range.

Three studies reported the FEV₁, expressed either in litres/second (L/s) (Lewith 2002) or as the percentage of the predicted value (Qutubuddin 2019, Reilly 1994). A FEV₁ value lower than predicted generally indicates lung obstruction. The MCID in FEV₁ for asthma has not been established, but it has been noted that changes of 200 mL in adults (or 12%) may be considered clinically important (I16).

Results from one study (Lewith 2002) (202 participants) suggested participants in the placebo group had a greater improvement in FEV₁ (0.414 L/sec) at the end of treatment (16 weeks) than those in the homeopathy group (0.136 L/sec) (MD 0.28; 95% CI -0.14, 0.69; $p = 0.19$). Pooled results from the other 2 studies (Qutubuddin 2019, Reilly 1994) (total 168 participants) suggested there is no important difference between the treatment groups (MD -1.67, 95% CI -19.71, 16.37, $p = 0.86$; $I^2 = 90\%$), but there was statistically high heterogeneity. (GRADE: Very Low).

One study (White 2003) reported the peak expiratory flow rate (PEFR) at the end of treatment (12 months). The PEFR is a measure of maximum speed of air forcefully expelled from the lungs, performed on a spirometer or similar device (I18). An MCID for the PEFR is not established, with peak flow measurements during steady state for each patient preferable to using published norms (which are dependent on the participant's age, weight, height etc.) (I19).

Results from this study (total 89 participants) were reported as a binary measure, based on the number of participants who reported less than a 15% improvement in PEFR (or $\geq 15\%$). No difference in PEFR improvement between the homeopathy and placebo groups was observed (RR 1.14; 95% CI 0.86, 1.53; $p = 0.36$).

Quality of life

Two studies (total 291 participants) reported quality of life measured using the Childhood Asthma Questionnaire (CAQ) – active quality of living domain or the Asthma Bother Profile at end of treatment (range 16 weeks to 12 months) (Lewith 2002, White 2003).

The CAQ comprises up to 31 questions for children to answer and up to 6 questions for parents, depending on the age of the child. The CAQ - active living domain measures how children feel during their daily activities. A pooled standardised score ranging from 0-100 (higher scores indicated a better quality of life) across all age groups was reported, with a clinically relevant difference considered a 7-point improvement on this scale (I08). Results from one study (White 2003) suggested there was no difference in quality of life between the homeopathy and placebo groups (MD 1.32; 95% CI -3.98, 6.62, $p = 0.63$). Changes from baseline to 12 months in other subscales of CAQ (severity, distress, passive quality of living) were reported per age group (not aggregated). Authors noted there was evidence of a general reduction (improvement) in the scores, but the differential treatment effect size was small.

The ABP consists of 23 items and was developed to assess how much asthma bothers patients (120). It has been shown to be a valid measure of health status (121). A total score is obtained by summing the scores of the items in sections 2, 3, and 4; with a maximum score of 75 indicating maximum bother and the minimum score of 0 indicating no bother. Results from one study (Lewith 2002) suggested there was no important difference between treatment groups at the end of treatment (16 weeks) (MD -0.03; 95% CI -0.20, 0.15; $p = 0.76$).

Taken together the results suggested there is no important difference between homeopathy or placebo on quality of life (SMD 0.00; 95% CI -0.23, 0.23; $p = 0.98$). (*GRADE: Low*).

Medication use

Two studies (291 participants) reported medication use by measuring the frequency of participants' inhaler use throughout the study period (range 16 weeks to 12 months) (Lewith 2002, White 2003).

One study (White 2003) reported the proportion of participants who had either increased, reduced or no change in their inhaler use; with no difference between the homeopathy group and the placebo group observed (RR 0.96, 95% CI: 0.68, 1.35; $p = 0.79$). (*GRADE: Low*).

Results from one study (Lewith 2002) were not able to be included in the meta-analysis, as the study authors did not report usable data (no means, SD, etc.). The study authors reported no evidence of a significant change in inhaler use in either the homeopathy or the placebo group.

D5.1.3.2 Secondary Comparison (vs 'inactive' control)

One RCT (Topcu 2010) and one Quasi RCT (Thompson 2008) comparing homeopathy with a control (no intervention) in people with asthma were eligible for this comparison. Each study contributed data to 3 outcomes considered critical for this review.

There were no ongoing studies and one study awaiting classification (total 60 participants) that compared homeopathy to a control that was published in language other than English.

No sensitivity analysis was performed examining the impact of studies at high risk of bias as no studies contributing data were at high risk of bias.

Asthma symptoms

Two studies (total 86 participants) measured asthma symptoms using the Asthma Control Questionnaire (ACQ) (range 16 to 26 weeks) (Thompson 2008, Topcu 2010). The ACQ consists of 7 questions relating to asthma symptoms, activity limitation, use of medications and pulmonary function, measured on a 7-point scale (0 = totally controlled, 7 = extremely poorly controlled). A change in score of 0.5 or more is considered the MCID for this outcome measure (114).

Pooled results (total 86 participants) showed no important difference in ACQ scores between the homeopathy and control (no intervention) groups (SMD 0.21, 95% CI -0.21, 0.64; $p = 0.33$; $I^2 = 0\%$) (*GRADE: Low*).

Pulmonary function

Two studies (total 86 participants) measured pulmonary function and reported the morning and evening peak expiratory flow rate (PEFR) at the end of treatment (range 16 to 52 weeks) (Thompson 2008, Topcu 2010).

The PEFR is a measure of maximum speed of air forcefully expelled from the lungs performed on a spirometer or similar device (118). An MCID for the PEFR is not established, with peak flow measurements during steady state for each patient preferable to using published norms (which are dependent on the participant's age, weight, height etc.) (119).

Results from one study (Thompson 2008) (total 35 participants) suggested an improvement in PEFR favouring the control (no intervention) group at 16 weeks, for both morning PEFR (MD 61.0; 95% CI 5.24, 116.76; $p = 0.03$) and evening PEFR (MD 70.0; 95% CI 16.15, 123.85; $p = 0.01$). (*GRADE: Very Low*)

Results from one study (Topcu 2010) (total 51 participants) were not included in the data synthesis, as the authors did not report usable data (means, SD, etc.). The study authors reported there was no significant difference in the absolute change from baseline between the homeopathy and placebo groups at 52 weeks (p -value not reported).

Quality of life

Two studies (total 86 participants) reported quality of life, measured with the Asthma Quality of Life Questionnaire (AQLQ) or the Paediatric AQLQ (PAQLQ) at end of treatment (range 16 to 52 weeks) (Thompson 2008, Topcu 2010). The AQLQ is a self-administered questionnaire consisting of 32 questions in 4 domains (symptoms, activity limitation, emotional function, and environmental stimuli). The PAQLQ contains 23 questions in three domains (symptoms, activity limitation and emotional function).

Participants are asked to think about how they have been during the previous week (or 2 weeks) and to respond to each question on a 7-point scale, from 1 (extremely bothered) to 7 (not bothered at all). The overall AQLQ or PAQLQ score is the mean of all responses, and the individual domain scores are the means of the items in those domains. A change in score of 0.5 is accepted as the MCID across each item of the AQLQ and PAQLQ (122).

Pooled results (total 89 participants) suggested there is no difference in quality of life between the homeopathy group and the control group at the end of treatment for any domain: (*GRADE: Low*)

- symptoms: MD -0.05; 95% CI -0.47, 0.36; $p = 0.80$; $I^2 = 0\%$
- activity limitation: MD -0.10; 95% CI -0.54, 0.34; $p = 0.65$; $I^2 = 0\%$
- emotional function MD 0.10; 95% CI -0.27, 0.47; $p = 0.60$; $I^2 = 0\%$
- environmental stimuli MD 0.10; 95% CI -0.45, 0.65, $p = 0.72$; $I^2 =$ not applicable

Hospitalisation

One study (total 39 participants) reported the number of participants who required inpatient care during the study (16 weeks). The results suggested that there was no important difference between the homeopathy or control groups (RR 1.06; 95% CI 0.17, 6.70; $p = 0.95$). (*GRADE: Low*)

Medication use

Two studies (total 86 participants) reported medication use based on the frequency of inhaler use at end of treatment (range 16 to 52 weeks) (Thompson 2008, Topcu 2010).

One study (Thompson 2008) reported the mean number of doses required per week, with the results suggesting there was no important difference between the homeopathy or control groups (MD -2.20; 95% CI -49.01, 44.61; $p = 0.93$). (*GRADE: Low*)

The other study (Topcu 2010) reported the median number of puffs per day, but the data were incomplete and not able to be included in the data synthesis. The study authors reported that there was no important difference between the homeopathy or control groups.

D5.1.3.3 Tertiary Comparison (vs other)

One RCT (Topcu 2010) (51 participants) comparing homeopathy with 'other' interventions in people with asthma was eligible for this comparison and contributed data to 3 outcomes. Data from these studies are presented in Appendix F2 Supplementary outcome data.

D6 Diseases of the digestive system

D6.1 Diarrhoea

D6.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-23. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-23 Overview of PICO criteria of included studies: Diarrhoea

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Jacobs 1993 (123)	Quasi-RCT	Diarrhoea, acute childhood	Individual homeopathy, oral	Placebo	Oral rehydration treatment	Symptom duration Symptom severity
Jacobs 2000 (124)	RCT	Diarrhoea, acute childhood	Individual homeopathy, oral	Placebo	Oral rehydration treatment	Symptom duration Symptom severity
Jacobs 2006 (125)	RCT	Diarrhoea, acute childhood	Non-individualised, oral combination	Placebo	Oral rehydration treatment	Symptom duration Symptom severity
Patel 2010 (126)	Quasi-RCT	Diarrhoea, acute childhood	Individual homeopathy, oral (acute) OR Individual homeopathy, oral (acute then constitutional remedy)	Placebo	Oral or IV rehydration treatment	Symptom severity
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus 'other' intervention**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial; IV, intravenous

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D6.1.2 Risk of bias per item

The risk of bias for each item in the included studies for diarrhoea is described below and shown graphically in Figure D-12 (details are provided in Appendix E).

Bias arising from the randomisation process

Two studies (Jacobs 1993, Jacobs 2000) were assessed to have some concerns for this domain. The details of allocation sequence randomisation, concealment, and baseline demographics were not provided for Jacobs 1993, although the study was described as a double-blind randomised control trial. Jacobs 2000 specified the method for generating the randomisation sequence and allocation concealment, however there was a slight imbalance in baseline characteristics, with the placebo group being younger, shorter, and lighter. This was not considered to be related to the randomisation process. One study (Jacobs 2006) was assessed as low risk of bias for this domain as the method of generating the randomisation sequence was specified, allocation sequence concealed, and baseline characteristics were well balanced. One study (Patel 2010) was assessed as high risk of bias as details of the allocation sequence and information on baseline characteristics were not provided.

Bias due to deviations from intended interventions

Three studies (Jacobs 1993, Jacobs 2000, Jacobs 2006) were assessed as low risk of bias for this domain as any deviations from the intended intervention reflected usual practice with their impact on the outcomes expected to be slight and analysis appropriate. One study (Patel 2010) was assessed as high risk of bias for this domain as the method of analysis was inappropriate and likely excluded eligible participants which could have biased the results.

Bias due to missing outcome data

Three studies (Jacobs 1993, Jacobs 2000, Jacobs 2006) were assessed as low risk of bias for this domain as data were available for nearly all participants, the analysis addressed missing data, or the proportions of and reasons for missing participants were similar across intervention groups.

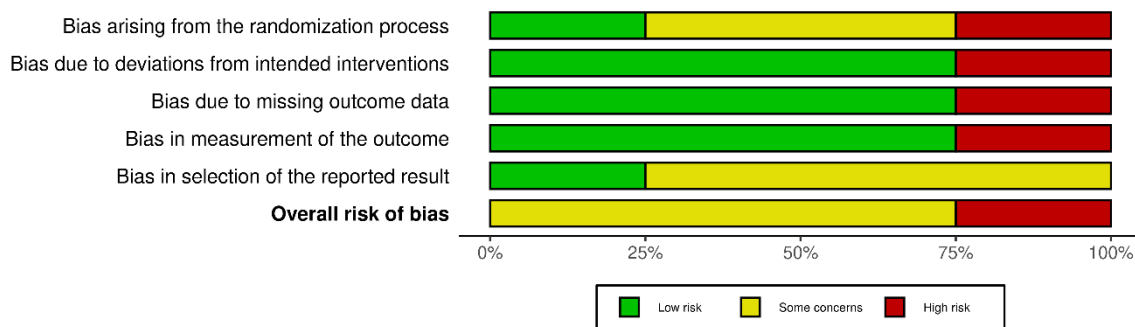
One study (Patel 2010) was assessed as high risk of bias for this domain due to the large proportion of missing data which was related to the true value of the outcome. The study reported that 24 cases were withdrawn from the study due to clinical worsening requiring hospitalisation.

Bias in measurement of the outcome

Three studies (Jacobs 1993, Jacobs 2000, Jacobs 2006) were assessed as low risk of bias for this domain as the outcome assessors were unaware of the intervention received by study participants. One study (Patel 2010) was assessed as high risk of bias for this domain as there was unclear blinding meaning participants and researchers measuring the outcome may have been influenced by knowledge of the intervention being received.

Bias in selection of the reported result

Three studies (Jacobs 1993, Jacobs 2006, Patel 2010) were assessed to have some concerns for this domain as no pre-specified analysis plans were provided. One study (Jacobs 2000) was assessed as low risk of bias for this domain as there was clear evidence that all reported results corresponded to all intended outcomes, analyses, and sub-cohorts.

Figure D-12 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs –diarrhoea

D6.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in digestive disorders (infantile diarrhoea) are listed in Table D-24.

Table D-24 Outcomes considered by the NTWC to be critical or important for decision-making: Diarrhoea

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Jacobs 1993	Jacobs 2000	Jacobs 2006	Patel 2010
Symptom severity	Number of stools, stool frequency	Critical	Yes	✓	✓	✓	✓ ^ ^
Symptom duration	Duration of diarrhoea (days)	Critical	Yes	✓	✓ ^	✓	--
Hospitalisations	--	Critical	No	--	--	--	--

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

^ Reported as probability of being diarrhoea free

^^ Reported as proportion of people with reduced symptom severity

D6.1.3.1 Primary Comparison (vs placebo)

Two RCTs (Jacobs 2000, Jacobs 2006) and 2 quasi-RCTs (Jacobs 1993, Patel 2010) (total 748 participants) comparing homeopathy with placebo in children with diarrhoea were eligible for this comparison. The studies contributed data to 2 of the 3 outcome domains considered critical for this review. There was no data available for one of the 3 outcome domains considered critical, probably because the outcome was not assessed in the studies.

There was one study awaiting classification (81 participants) that compared homeopathy with placebo in children with diarrhoea. This study could have contributed data to at least one outcome considered critical for this review, however its full text was not able to be retrieved.

Symptom severity

Three studies (total 448 participants) reported symptom severity based on the mean number of loose stools per day (Jacobs 1993, Jacobs 2000, Jacobs 2006). Two studies did not provide complete data, so could not be included in the meta-analysis.

One study (Jacobs 1993) suggested there was no difference in symptoms severity between the homeopathy and placebo groups ($p = 0.57$). One study (Jacobs 2000) suggested improvement in symptom severity favoured the homeopathy group ($p = 0.023$). The remaining study (Jacobs 2006) (292 participants) also reported improvement in symptom severity did not differ between homeopathy and placebo groups (MD -0.02 ; 95% CI $-0.76, 0.36$; $p = 0.48$) (*GRADE: Low*).

No sensitivity analysis was performed examining the impact of studies at high risk of bias as no studies contributing data were at high risk of bias.

One quasi-RCT (total 300 participants) (Patel 2010) reported the proportion of participants with reduced symptom severity based on a clinician grading scale (including vomiting, stool frequency, stool quantity). The trialists reported change in diarrhoea severity as either “aggravation”, “amelioration”, or “status quo” based on changes in grading of diarrhoea 24-hours after treatment initiation. The results suggested significantly more children experienced amelioration of symptoms in the homeopathy group compared with the placebo group (RR 3.67; 95% CI 2.39, 5.64; $p < 0.00001$). (*GRADE: Very Low*)

No sensitivity analysis was performed examining the impact of studies at high risk of bias as all studies contributing data were at high risk of bias.

Symptom duration

Three studies (total 448 participants) reported symptom duration based on participant recall and symptom diaries. Two studies (Jacobs 1993, Jacobs 2006) reported symptom duration in days. One study (Jacobs 2000) reported symptom duration as the probability of being symptom free after 5 days.

All 3 studies provided data in a format that was not able to be synthesised for meta-analysis. Two of the studies (Jacobs 1993, Jacobs 2006) (total 325 participants) reported no significant difference between homeopathy and placebo groups; with one study (Jacob 2006) also noting that children in the treatment group were equally likely to resolve their diarrhea at any given time during follow-up as children in the placebo group (HR 1.02; 95% CI 0.79, 1.33). One study (Jacobs 2000) (123 participants) suggested it was more likely for the homeopathy group to be symptom free after 5 days (42.1%) compared to the placebo group (60.5%) (*GRADE: Low*).

No sensitivity analysis was performed examining the impact of studies at high risk of bias as no studies contributing data were at high risk of bias.

D6.1.3.2 Secondary Comparison (vs ‘inactive’ control)

There were no RCTs comparing homeopathy with inactive control in digestive disorders (infantile diarrhoea).

D6.1.3.3 Tertiary Comparison (vs other)

There were no RCTs comparing homeopathy with ‘other’ interventions in digestive disorders (infantile diarrhoea).

D6.2 Infantile colic and other digestive disorders

D6.2.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-25. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-25 Overview of PICO criteria of included studies: Infantile colic and other digestive disorders

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Dossett 2015 (127)	RCT	Gastroesophageal reflux disease (GERD)	Non-individualised homeopathy, oral	Placebo	Standard length interview OR expanded length interview	Symptom severity Quality of life
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
Paterson 2003 (128)	RCT	Dyspepsia	Individualised homeopathy, oral	Inactive control (usual care) OR Acupuncture [^]	None reported	Disease severity/symptoms Quality of life
Homeopathy versus 'other' intervention**						
Raak 2019 (129)	RCT	Infantile colic	Non-individualised homeopathy, oral	Simethicone, oral	None reported	Symptom duration Symptom frequency

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

[^] Patterson 2003 included 3 treatment arms (individualised homeopathy, normal general practitioner care and acupuncture). Homeopathy compared to normal general practitioner care is considered for the comparison versus inactive control. Homeopathy compared to acupuncture is considered for the comparison versus 'other' intervention.

D6.2.2 Risk of bias per item

The risk of bias for each item in the included studies for infantile colic and other digestive disorders are described below and shown graphically in Figure D-13 (details are provided in Appendix E).

Bias arising from the randomisation process

One study (Dossett 2015) was assessed to have some concerns of bias for this domain. The details of allocation sequence randomisation and concealment were provided, however there were some differences in baseline characteristics between groups. This was not considered to be related to the randomisation process.

Two studies (Paterson 2003, Raak 2019) were assessed to have low risk of bias for this domain as the details of allocation sequence and concealment were provided and baseline characteristics were comparable between groups.

Bias due to deviations from intended interventions

All 3 studies (Dossett 2015, Paterson 2003, Raak 2019) were assessed to have low risk of bias for this domain as any deviations from the intended intervention reflected usual practice with their impact on the outcomes expected to be slight and the analysis used appropriate.

Bias due to missing outcome data

All 3 studies (Dossett 2015, Paterson 2003, Raak 2019) were assessed to have low risk of bias for this domain as data were available for nearly all participants.

Bias in measurement of the outcome

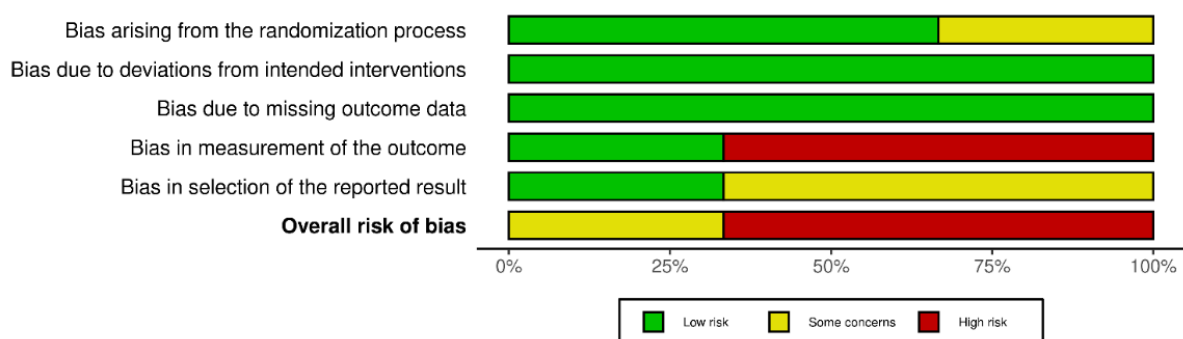
Two studies (Paterson 2003, Raak 2019) were assessed to have high risk of bias for this domain. One study (Paterson 2003) used an inappropriate subjective measure for a primary outcome which differed between intervention groups. In one study (Raak 2019), both investigators and participants knew which intervention they received which may have influenced the measurement of the outcome.

One study (Dossett 2015) was assessed to have low risk of bias for this domain as methods of outcome assessment were appropriate and comparable across treatment groups.

Bias in selection of the reported result

Two studies (Paterson 2003, Raak 2019) were assessed to have some concerns for this domain as no information of a pre-specified analysis plan was provided. One study (Dossett 2015) was assessed to have low risk of bias for this domain there as all reported results correspond to all intended outcomes and analyses.

Figure D-13 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – infantile colic and other digestive disorders



D6.2.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with infantile colic and other digestive disorders are listed in Table D-26

These disease-specific outcomes measures were prioritised with infantile colic regarded as the primary condition of interest. Studies in other digestive disorders were also to be included, but the prioritised measures did not correlate. For these studies the reviewers reverted to the outcome domains the NWTC had originally proposed as being critical or important for decision-making (see Appendix A6.2; Table A-11), with evidence in two domains identified for inclusion in the synthesis (disease severity/symptoms and quality of life). One study (Dossett 2015) provided data for the main comparison and one study (Paterson 2003) provided data for Secondary Comparison.

Table D-26 Outcomes considered by the NTWC to be critical or important for decision-making: Infantile colic and other digestive disorders

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Paterson 2003	Dossett 2015	Raak 2019
Symptom duration	Duration of crying	Critical	No	--	--	--^
Symptom frequency	Frequency of crying (per 24 hours)	Critical	No	--	--	--^
Complaint score	Response rate on composite symptom measure	Critical	No	--	--	--^
Objective symptom score	Response rate on composite symptom measure	Critical	No	--	--	--^

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

^ Study measured outcome domain however is not available for main comparison (homeopathy versus 'other' intervention).

D6.2.3.1 Primary Comparison (vs placebo)

One RCT (Dossett 2015) (24 participants) comparing homeopathy with placebo in gastroesophageal reflux disease (GERD) was eligible for this comparison. The study contributed to 2 outcome domains considered critical or important for decision-making.

There was one study awaiting classification (40 participants) that compared homeopathy with placebo in GERD. This study could have contributed data to the outcome domains critical for this review, however, the article was published in a language other than English.

No sensitivity analysis was performed examining the impact of studies at high risk of bias as no studies contributing data were at high risk of bias.

Symptom severity

One RCT (Dossett 2015) reported symptom severity based on participant symptom diaries at the end of treatment (2 weeks). The severity of nine different GERD and dyspepsia-related symptoms (daytime heartburn, nighttime heartburn, acid regurgitation, upper abdominal pain, fullness after eating, early satiety, flatulence, belching, and nausea) were recorded daily by participants according to a 5-point scale (none, mild, moderate, severe, or very severe) (higher is worse). The total daily GERD symptom severity was based on the sum of scores assessing severity of daytime heartburn, nighttime heartburn, and acid reflux (range 0 to 12). The total daily dyspepsia symptom severity was based on the sum of scores of the remaining symptoms (range 0 to 24).

The results from one study (total 24 participants) suggested participants in the placebo group had better symptom control, but there was no difference in symptom severity between homeopathy and placebo groups for symptoms attributed to GERD (MD 1.10; 95% CI -0.56, 2.76; $p = 0.19$) or symptoms of dyspepsia (MD 0.30; 95% CI -1.86, 2.46; $p = 0.79$) (GRADE: Very Low).

Quality of life

One RCT (Dossett 2015) reported quality of life using the GERD-Health-Related Quality of Life Instrument (GERD-HRQL score) at the end of treatment (2 weeks). The GERD-HRQL is a reliable and valid measurement of GERD symptoms, including heartburn and regurgitation in response to treatment (130). Scores range from 0 to 75, with a greater score indicating worse symptoms. No MCID for the GERD-HRQL was found.

The results from one study (total 24 participants) suggested participants in the placebo group had better quality of life, but the difference between homeopathy and placebo groups was not significant (MD 4.35; 95% CI -0.51, 9.21; $p = 0.08$) (GRADE: Very Low).

D6.2.3.2 Secondary Comparison (vs 'inactive' control)

One study (Paterson 2003) (60 participants) comparing homeopathy with usual care was eligible for this comparison. The study contributed to 2 outcome domains considered critical or important for decision-making in homeopathy.

Symptom severity

One RCT (Paterson 2003) reported symptom severity using the Measure Yourself Medical Outcome Profile (MYMOP). The MYMOP is an individualised measure that asks patients to nominate and report on the effect of 2 disease-related symptoms (131). The overall score is calculated as the average of item scores, with a higher score indicating worse symptoms. The outcome was measured as change from baseline (6 weeks).

Results from one study (40 participants) suggested there was no difference in improvement of symptom severity between the homeopathy and the control groups (MD -0.09; 95% CI -1.08, 0.90; $p = 0.86$). (GRADE: Very Low).

Quality of life

One RCT (Paterson 2003) reported quality of life using the SF-36 health survey. However, the study authors only provided data for participants baseline SF-36 healthy survey scores and therefore the data could not be included for this comparison.

D6.2.3.3 Tertiary Comparison (vs other)

Two RCTs (Paterson 2003 and Raak 2019) comparing homeopathy with 'other' interventions were eligible for this comparison. One study (Paterson 2003) compared homeopathy with acupuncture in people living with GERD and contributed data for 2 outcomes considered critical or important for decision-making in homeopathy. One study (Raak 2019) comparing homeopathy with simethicone for infantile colic contributed data for 4 of the 4 disease-specific outcomes highlighted by the NWTC. Data from these studies are presented in Appendix F2 Supplementary outcome data.

D6.3 Irritable bowel syndrome

D6.3.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-27. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-27 Overview of PICO criteria of included studies: Irritable bowel syndrome

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
No studies found						
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
Peckham 2012 (132, 133)	Quasi RCT	Irritable bowel syndrome	Individualised homeopathy, 5 x 1-hour sessions over 6 months	No intervention OR Supportive listening [^]	Usual care	IBS Symptom Severity Anxiety Depression HRQoL
Homeopathy versus 'other' intervention**						
No studies found						

Abbreviations: HRQoL, health-related quality of life; RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

[^] Study included 3 groups. The inactive control (no intervention) is considered in the evidence synthesis.

D6.3.2 Risk of bias per item

The risk of bias for each item in the included studies for IBS is described below and shown graphically in Figure D-14 (details are provided in Appendix E).

Bias arising from the randomisation process

Peckham 2012 was judged to have some concerns for this domain, due to the quasi randomisation process.

Bias due to deviations from intended interventions

Peckham 2012 was considered at low risk of bias for this domain, as there were no deviations from the intended intervention considered to have arisen due to the trial context, and an appropriate method of analysis was used.

Bias due to missing outcome data

Peckham 2012 was assessed to have some concerns for this domain, as 12.8% of participants had missing follow up data, with no adjustment for this missingness being presented. Reasons for drop out are not reported, making it difficult to assess whether missingness was due to the true value of the outcome.

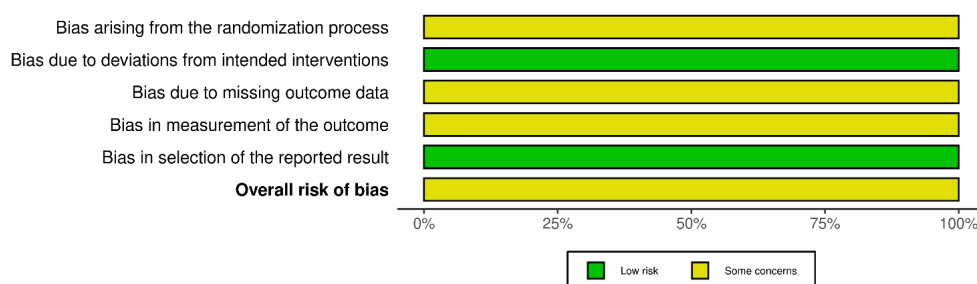
Bias in measurement of the outcome

Peckham 2012 was judged to have some concerns for this domain, due to the self-reporting of outcomes by non-blinded participants. There was no indication that participants were likely to be biased in their reporting, however it was considered possible.

Bias in selection of the reported result

Peckham 2012 was considered at low risk of bias for this domain, as the presented results aligned with the pre-specified analysis plan. 52-week results were not available despite being pre-specified, however a justification for this was provided and considered reasonable.

Figure D-14 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – irritable bowel syndrome



D6.3.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with IBS are listed in Table D-28.

Table D-28 Outcomes considered by the NTWC to be critical or important for decision-making: Irritable bowel syndrome

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Peckham 2012
Symptom severity	IBS Severity Scoring System (or other validated measure)	Critical	Yes	✓
Pain	Abdominal Pain Index, VAS (or other validated measure)	Critical	No	--
Quality of life	IBS Quality of Life (or other validated measure)	Critical	Yes	✓

Abbreviations: IBS, irritable bowel syndrome; VAS, visual analogue scale

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

D6.3.3.1 Primary Comparison (vs placebo)

There were no studies found that compared homeopathy with placebo in people with IBS.

There were 3 studies awaiting classification (160+ participants) that compared homeopathy with placebo: 2 studies were published in a language other than English (published prior to 1980) and one study (total 60 participants) had results that were published after the literature search date (134). This study could have contributed data to 2 outcomes considered critical or important by the NTWC.

There are 2 ongoing studies (planned 161 participants) comparing homeopathy to placebo (not yet complete) that will contribute data to 2 outcomes considered critical or important by the NTWC.

D6.3.3.2 Secondary Comparison (vs 'inactive' control)

There was one quasi RCT (Peckham 2012) comparing homeopathy with inactive control (no intervention) in people with IBS that was eligible for this comparison and contributed data for 2 outcomes considered critical or important for this review.

There were no studies awaiting classification or ongoing that were eligible for this comparison.

No sensitivity analyses were performed examining the impact of studies at high risk of bias as only one study contributed data that was not at high risk of bias.

Symptom severity

One study (76 participants) reported symptom severity measured using the IBS Symptom Severity Scale (IBS-SSS) at the end of treatment (26 weeks) (Peckham 2012).

The IBS-SSS is a disease-specific questionnaire incorporating measures of pain, distension, bowel dysfunction and quality of life/global wellbeing (135). Scores range from 1 to 500, with higher scores indicating more severe IBS symptoms. A 50-point change in scores is considered clinically relevant (136). Results suggested there was no difference in symptom severity in the homeopathy group compared to the control group (MD -26.86; 95% CI -88.59, 34.87; $p = 0.39$) (*GRADE: Low*).

Quality of life

One study (76 participants) reported quality of life measured using the EQ-5D-3L at the end of treatment (26 weeks) (Peckham 2012).

The EQ-5D-3L is a standardised health-related quality of life measure which measures an individual's health state on 5 scales: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has three response levels of severity: no problems, some problems, extreme problems. The dimensional scores are subsequently used to derive a utility score. The MCID for the EQ-5D-3L utility score in people IBS is not established but is estimated to be 0.08 in people with inflammatory bowel disease (137). The study did not report a calculated utility score therefore the results could not be included in the data synthesis.

The EQ-5D-3L also includes a visual analogue scale of self-rated quality of life on a scale from 0 (worst health you can imagine) to 100 (best health you can imagine). The MCID of EQ-5D-3L VAS is not established, but is estimated to be 10 points in people with inflammatory bowel disease (137).

Results for the EQ-5D VAS showed no difference in quality of life in the homeopathy group compared to the control group (MD -5.66; 95% CI -16.00, 4.69; $p = 0.28$) (*GRADE: Low*).

D6.3.3.3 Tertiary Comparison (vs other)

One quasi RCT (Peckham 2012) comparing homeopathy with 'other' interventions in people with IBS was eligible for this comparison and contributed data for 2 outcomes. Data from these studies are presented in Appendix F2 Supplementary outcome data.

D7 Diseases of the skin

D7.1 Psoriasis

D7.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-29. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-29 Overview of PICO criteria of included studies: Psoriasis

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Wiesenauer 1992 (138)	Quasi RCT	Psoriasis	Non-individualised homeopathy (<i>M. aquifolium</i> 10%), topical	Placebo	None reported	Disease severity
Bernstein 2006 (139)	Quasi RCT	Psoriasis	Non-individualised homeopathy (<i>M. aquifolium</i> 10%), topical	Placebo	None reported	Disease severity Quality of life
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found.						
Homeopathy versus other 'active' intervention**						
No studies found.						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D7.1.2 Risk of bias per item

The risk of bias for each item in the included studies for psoriasis is described below and shown graphically in Figure D-15 (details are provided in Appendix E).

Bias arising from the randomisation process

Two studies (Wiesenauer 1992 and Bernstein 2006) were assessed to be at low risk of bias for this domain. The details of the randomisation process and allocation concealment were not provided for Wiesenauer 1992 and Bernstein 2006, however, both studies are described as double-blind randomised control trials.

Baseline differences were not able to be compared in one study (Wiesenauer 1992) due to the intraindividual application of both the placebo and homeopathy to each participant. Baseline characteristics were well balanced between groups in Bernstein 2006.

Bias due to deviations from intended interventions

Two studies (Wiesenauer 1992 and Bernstein 2006) were assessed to be at low risk of bias for this domain because both performed an intent-to-treat analysis with participants analysed in the group to which they were randomised.

Bias due to missing outcome data

One study (Wiesenauer 1992) was assessed as low risk of bias for this domain as data was available for nearly all participants with less than 5% missingness.

One study (Bernstein 2006) was assessed as high risk of bias for this domain, as there was greater than 5% missingness in outcome data (29 discontinued participants from 200 randomised participants). Of the discontinued participants, there was unequal distribution of missing outcome data between groups, favouring homeopathy (3 discontinued participants in homeopathy group and 26 discontinued in placebo group). Additionally, the authors inappropriately imputed the worst possible score for all discontinued participants with missing outcome data in the intent-to-treat analysis. As there was greater missingness in the placebo group, such imputation favoured homeopathy.

Bias in measurement of the outcome

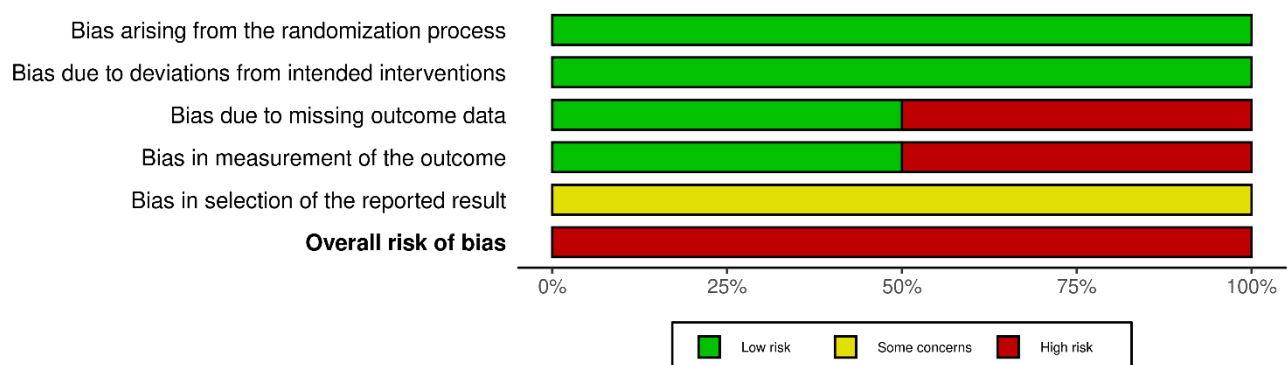
One study (Wiesenauer 1992) was assessed as high risk of bias for this domain. Although measurement of the outcome did not differ between groups, the main effect measure relied on self-assessment, was crude, and unlikely to be sensitive to intervention effects. Additionally, as the details of allocation sequence and concealment were not provided, measurement of the outcome may have been influenced by knowledge of the intervention being received.

One study (Bernstein 2006) was assessed as low risk of bias for this domain as the outcome measures used were appropriate, objective, consistent between groups, and unlikely to be influenced by knowledge of the intervention received by study participants.

Bias in selection of the reported result

Both studies (Wiesenauer 1992, Bernstein 2006) were assessed to have some risk of bias for this domain as no pre-specified analysis plans were provided.

Figure D-15 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – psoriasis



D7.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with psoriasis are listed in Table D-30.

Table D-30 Outcomes considered by the NTWC to be critical or important for decision-making: Psoriasis

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Wiesenauer 1992	Bernstein 2006
Disease severity/symptoms	Psoriasis Area Severity Index (PASI) or clinician assessed	Critical	Yes	✓	✓
Quality of life	Quality of life index (QLI) questionnaire	Critical	Yes	--	✓
Itching	Numeric rating scale	Critical	No	--	--
Skin condition	Numeric rating scale	Critical	No	--	--
Medication use	Steroid creams or ointments	Critical	No	--	--

Abbreviations: PASI, Psoriasis Area Severity Index; QLI, Quality of life index

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

--No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

D7.1.3.1 Primary Comparison (vs placebo)

Two quasi-RCTs (Bernstein 2006, Wiesenauer 1992) comparing homeopathy with placebo were eligible for this comparison. The studies contributed data to 2 of the 5 outcome domains considered critical for this review.

There were no studies awaiting classification and one ongoing study (target 60 participants) that compared homeopathy with placebo in people with psoriasis. The study is not yet complete but could contribute to at least 2 outcomes considered critical for this review (see Appendix C4).

Disease severity

One study (total 200 participants) reported psoriasis severity measured using the Psoriasis Area Severity Index (PASI) at the end of treatment (12 weeks) (Bernstein 2006).

The PASI is the gold standard in the assessment of the extent and severity of psoriasis, with a higher score indicating greater severity (140). A representative area of psoriasis is selected using a 4.0 cm² template. The intensity of redness, thickness, and scaling of the psoriasis is assessed as none (0), mild (1), moderate (2), severe (3), or very severe (4). The percent amount of skin in the area affected is also recorded (0 to 100%), with a total score calculated as a percentage of the above score⁶. A 50% reduction in the PASI score is considered clinically meaningful (141).

Results from one study (200 participants) suggested an improvement in PASI score that favoured homeopathy when compared with placebo (MD -3.30; 95% -4.48, -2.12; $p < 0.00001$); but the size of improvement (< 30%) is not clinically important (*GRADE: Low*).

⁶ Bernstein 2006 measures and reports a PASI score for one body area, whereas current clinical approach is to assess 4 body regions for a maximum score of 72.

One study (Wiesenauer 1992) (total 80 participants) reported the proportion of participants with unchanged, improved, and disappeared psoriasis symptoms as per independent self-assessment and physician assessment at the end-of treatment (median 4 weeks, IQR 21 to 49 weeks).

Based on physician assessment, the results suggested more participants had improvement or resolution of psoriasis symptoms in the homeopathy group (72.5%) compared with the placebo group (45%) (RR 1.61; 95% CI 1.09, 2.38; $p = 0.02$). (*GRADE: Very Low*).

Quality of life

One study (Bernstein 2006) reported quality of life measured using the Quality of Life Index (QLI) questionnaire at the end of treatment (12 weeks). The questionnaire assesses quality of life in the previous month, with a higher score indicating poorer quality of life. Twelve questions are quantified from 0 (not at all) to 10 (very much), with a maximum possible score of 120. An MCID for the QLI was not found.

The results from one study (200 participants) suggested an improvement from baseline in quality of life favouring homeopathy when compared with placebo (MD -10.40; 95% CI -17.59, -3.21; $p = 0.005$). (*GRADE: Very Low*).

D7.1.3.2 Secondary Comparison (vs 'inactive' control)

There were no eligible RCTs comparing homeopathy with inactive control in people with psoriasis. There is one study awaiting classification (total 36 participants) that was eligible for this comparison that could have contributed data to at least one critical or important outcome. The study was published in a language other than English (see Appendix C4).

D7.1.3.3 Tertiary Comparison (vs other)

There were no eligible RCTs comparing homeopathy with 'other' interventions in psoriasis.

D8 Diseases of the musculoskeletal system or connective tissue

D8.1 Arthropathies

D8.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-31. The table is ordered by sub-populations and then alphabetically. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-31 Overview of PICO criteria of included studies: Arthropathies

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Koley 2015 (142)	RCT	Osteoarthritis (knee)	Individualised homeopathy, oral	Placebo	None reported Consultation likely for both groups	Pain Stiffness Physical function
Shipley 1983 (143)	Quasi RCT	Osteoarthritis (hip or knee)	Non-individualised homeopathy, oral	Placebo	Paracetamol permitted	Pain
Brien 2004 (144-146)	RCT	Rheumatoid arthritis	Individualised homeopathy (consultation) OR Non-individualised homeopathy (with or without consultation)	Placebo (with or without consultation)	None reported	Disease severity HRQoL Emotional function Pain Physical function/disability Disease biomarkers
Fisher 2001 (147)	Quasi RCT	Rheumatoid arthritis	Individualised homeopathy, oral	Placebo	DMARDs or NSAIDs Both groups received consultation with homeopath	Pain Tenderness Stiffness Disease biomarkers
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus 'other' intervention**						
Ibrahim 2015 (148)	RCT	Osteoarthritis (knee)	Non-individualised homeopathy, oral	Physiotherapy OR Acupuncture	Pharmacotherapy (NSAID and paracetamol)	Pain Stiffness Physical function HRQoL
Shealy 1998 (149)	Quasi RCT	Osteoarthritis (knee)	Non-individualised homeopathy, oral	Pharmacotherapy (paracetamol)	None reported	Pain
Strosser 2000 (150)	Quasi RCT	Osteoarthritis (knee)	Non-individualised homeopathy, oral	Pharmacotherapy (diclofenac)	None reported	Disease severity
Van Haselen 2000 (151)	RCT	Osteoarthritis (knee)	Non-individualised homeopathy, topical	Pharmacotherapy (piroxicam gel)	Paracetamol permitted	Pain Medication use HRQoL

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Widrig 2007 (152)	RCT	Osteoarthritis (hand)	Non-individualised homeopathy, topical	Pharmacotherapy (ibuprofen)	Paracetamol permitted	Pain Physical function Stiffness Medication use HRQoL
Khitrov 2009 (153)	Quasi RCT	Periarthritis of the shoulder	Non-individualised homeopathy, oral	Pharmacotherapy (NSAID)	None reported	Pain Range of movement Disease biomarkers HRQoL

Abbreviations: DMARD, disease modifying anti-rheumatic drug; HRQoL, health-related quality of life; NSAID, non-steroidal anti-inflammatory drug; RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D8.1.2 Risk of bias per item

The risk of bias for each item in the included studies for arthritis is described below and shown graphically in Figure D-16 (details are provided in Appendix E).

Bias arising from the randomisation process

Four studies (Fisher 2001, Khitrov 2009, Shipley 1983, Strosser 2000) were assessed at high risk of bias for this domain due to insufficient reporting of the randomisation method, allocation concealment and baseline participant characteristics. Two studies (Shealy 1998, Widrig 2007) were assessed to have some concerns due to lack of information regarding allocation concealment and not specifying the method of generating the randomisation sequence (Shealy 1998 only).

Three studies (Brien 2004, Koley 2015, van Haselen 2000) were judged at low risk of bias for this domain.

Bias due to deviations from intended interventions

Five studies (Fisher 2001, Khitrov 2009, Shealy 1998, Shipley 1983, Strosser 2000) were judged at high risk of bias for this domain. This was due to not reporting the method of analysis and not reporting the number of participants potentially analysed not in the group to which they were randomised (Khitrov 2009, Shealy 1998, Shipley 1983, Strosser 2000) or due to a per protocol method of analysis being used which excluded 12 out of 112 participants from the analysis (Fisher 2001).

Four studies (Brien 2004, Koley 2015, van Haselen 2000, Widrig 2007) were judged at low risk of bias for this domain.

Bias due to missing outcome data

Seven studies (Brien 2004, Fisher 2001, Khitrov 2009, Koley 2015, Shealy 1998, Shipley 1983, Strosser 2000) were assessed at high risk of bias for this domain. This was due to the high rate of dropout across the studies, which was judged as likely to be related to the true value of the outcome. A number of studies noted that participants dropped out due to aggravation or deterioration of symptoms (Koley 2015, Shipley 1983, Brien 2004) or change to conventional medication (Fisher 2001). Three studies (Khitrov 2009, Shealy 1998, Strosser 2000) did not provide information on the number of participants who dropped out.

Two studies (van Haselen 2000, Widrig 2007) were judged at low risk of bias for this domain.

Bias in measurement of the outcome

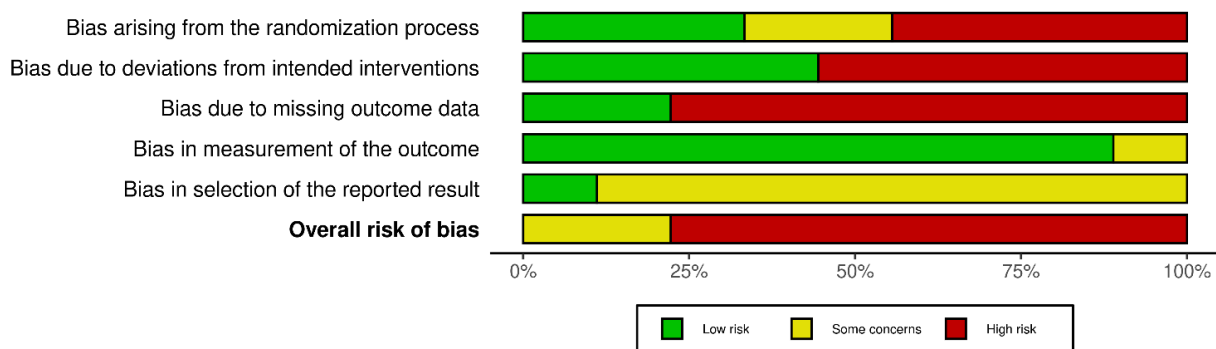
One study (Khitrov 2009) was assessed to have some concerns of bias for this domain, due to non-blinded participants who could plausibly be biased in their reporting of self-reported outcomes. The remaining studies were judged at low risk of bias for this domain as they reported that participants and trialists were blinded to intervention status.

Bias in selection of the reported result

Eight studies (Khitrov 2009, Koley 2015, Shealy 1998, Shipley 1983, Strosser 2000, van Haselen, Widrig 2007, Fisher 2001) were assessed to have some concerns for this domain due to the lack of a pre-specified analysis plan against which to compare the reported results. There was no evidence of inappropriate analysis or selective reporting of results.

One study (Brien 2004) was assessed at low risk of bias for this domain as reported results aligned with those in the trial protocol.

Figure D-16 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – arthropathies



D8.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with arthritis are listed in Table D-32.

Table D-32 Outcomes considered by the NTWC to be critical or important for decision-making: Arthropathies

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Brien 2004	Fisher 2001	Koley 2015	Shipley 1983
Pain intensity	VAS, WOMAC pain (or other validated measure)	Critical	Yes	✓*	✓^	✓	✓^
Physical functioning/ disability	WOMAC disability (or other validated measure)	Critical	Yes	✓	--	✓	--
Disease severity	WOMAC, ACR20 (or other validated measure)	Critical	Yes	✓	--	--	--
Quality of life	SF-36, EuroQoL (or other validated measure)	Critical	Yes	✓	--	--	--

Medication use	Paracetamol use	Critical	No	--	--	--	√^
Abbreviations: ACR, American College of Rheumatology Criteria 20% improvement; VAS, Visual Analogue Scale; WOMAC, Western Ontario and McMaster Universities Arthritis Index							
√ A study result is available for inclusion in the synthesis							
X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators							
-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results							
? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study							
* Current pain (VAS) nominated as the priority outcome. Weekly pain (symptom diary) also reported but not extracted here.							
^ Study does not report results prior to crossover so was not able to be used in the synthesis							

D8.1.3.1 Primary Comparison (vs placebo)

Two RCTs (Koley 2015, Brien 2004) and 2 quasi RCTs (Shiple 1983, Fisher 2001) comparing homeopathy with placebo were eligible for this comparison and contributed data to 5 out of 5 outcomes considered critical or important for this review.

There were 5 studies awaiting classification (total 375 participants) that could have contributed data to all outcomes. Of the 14 ongoing studies (total 1221 participants), 4 were considered likely to have been completed not reported because the *p* value, magnitude or direction of the results generated were considered unfavourable by the study investigators (see Appendix C6).

Pain intensity

Two studies (Brien 2004, Koley 2015) reported pain intensity measured with a VAS at the end of treatment (range: 2 to 24 weeks). In one study (Brien 2004), scores were reported as change from baseline to end of treatment (24 weeks). Results from 2 additional studies (Fisher 2001, Shiple 1983) (total participants unknown) were not included in the meta-analysis as both studies were crossover trials and study authors did not report results after the first treatment period. As per protocol, these results were not eligible for inclusion in the quantitative synthesis.

The VAS is subjective tool that can be used to measure a variety of outcomes. It is measured on a continuous scale (mm) from 0 (no pain) to 100 (worst imaginable pain), with higher scores indicating a higher intensity of pain. The median absolute MCID on a VAS scale in people with chronic pain is reported to be 20 mm (IQR 15–30) (154). In people with rheumatoid arthritis, the MCID has been estimated to be an 11.9-point change (155).

Pooled results from 2 studies (112 participants) suggested there is no difference in pain intensity for the homeopathy group compared to the placebo group (SMD 0.01; 95% CI –0.36, 0.39; *p* = 0.73; *I*² = 0%) (GRADE: Very Low).

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as both studies contributing data were judged to have high risk of bias.

Physical function/disability

Two studies (Brien 2004, Koley 2015) reported physical functioning/disability measured with the Health Assessment Questionnaire (HAQ) or a VAS at end of treatment (range: 2 to 24 weeks).

One study (Brien 2004) measured disability with the HAQ, reported as change from baseline to end of treatment (24 weeks). The HAQ is a measure of disability impact on daily life. It includes 20 questions across 8 sections: dressing, arising, eating, walking, hygiene, reach, grip and activities. The question with the highest score for each section becomes the score for that section. Scores for each section range from 0 to 3, with higher scores indicating greater difficulty. The total score is calculated as an average of each section. The MID in HAQ score has been shown to vary depending on baseline score, with the MID ranging from –0.03 to –0.14 for mild and more severe disease respectively (156). Results from one study (54 participants) suggested there is no difference in disability between the homeopathy and placebo group (SMD –0.17, 95% CI –0.71, 0.37; $p = 0.55$).

One study (Koley 2015) measured physical functioning with a VAS. The VAS is a subjective tool that can be used to measure a variety of outcomes. In Koley 2015, participants were asked to indicate their limitation of physical function on a 0-100mm VAS, where a higher score indicated more severe limitations. No MCID for physical function VAS in people with osteoarthritis was identified. Results from one study (60 participants) suggested there was no difference in physical function/disability between the homeopathy group and the placebo group (SMD 0.05, 95% CI –0.46, 0.56; $p = 0.85$).

Pooled results from 2 studies showed no difference in physical function/disability between homeopathy and placebo (SMD –0.05; 95% CI –0.42, 0.32; $p = 0.57$; $I^2 = 0\%$) (*GRADE: Very Low*).

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as both studies were judged to be at high risk of bias.

Disease severity

One study (Brien 2004) reported disease severity measured using the Disease Activity Score (DAS-28) at end of treatment (24 weeks). Results were reported as change from baseline to end of treatment (24 weeks).

The DAS-28 is a measure of RA severity using both clinical and laboratory data. It consists of both subjective (tender joint count and patient general assessment) and objective (swollen joint count and ESR or CRP) components. Scores range from 0 to 10, with higher scores indicating more severe disease activity. A decrease of 1.2 points to a total score of less than 3.2 is considered a 'good response', while a change of less than 0.6 is considered no response (157).

Results from one study (53 participants) suggested there was no improvement in disease severity in the homeopathy or placebo groups, with the change scores being around 0.6 points (MD 0.06; 95% CI –0.57, 0.68; $p = 0.85$) (*GRADE: Very Low*).

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data and it was judged to be at high risk of bias.

Quality of life

One study (Brien 2004) reported health related quality of life measured using a VAS at the end of treatment (24 weeks). Results were reported as change from baseline to end of treatment (24 weeks).

In Brien 2004, the VAS was a 100mm scale and participants were asked to rate their global assessment of health. No MCID for quality of life in people with RA was identified. Results from one study (54 participants) suggested there is no difference in health-related quality of life between the homeopathy and the placebo s (MD 4.96; 95% CI –8.78, 18.70; $p = 0.48$) (*GRADE: Very Low*).

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data and it was judged to have high risk of bias.

Medication use

One study (Shipley 1983) reported medication use measured using the number of paracetamol tablets returned at the completion of the study. This study reported pooled results across all arms after cross-over. Per protocol, these results were not eligible for inclusion in the quantitative synthesis.

D8.1.3.2 Secondary Comparison (vs 'inactive' control)

There were no eligible studies that compared homeopathy with inactive control in people with arthritis. There was one study awaiting classification (total participants unknown) that could have contributed data to the disease severity outcome.

D8.1.3.3 Tertiary Comparison (vs other)

There were 6 RCTs comparing homeopathy with 'other' interventions in people with arthritis that were eligible for this comparison and contributed data for 4 outcomes. Data from these studies are presented in Appendix F2 Supplementary outcome data.

D8.2 Back and/or neck pain

D8.2.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-33. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-33 Overview of PICO criteria of included studies: Back and/or neck pain

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Gupta 2020 (158)	RCT	Neck pain	Individualised homeopathy, oral	Placebo	None reported	Pain Stiffness/Mobility Quality of life
Morris 2016 (159)	RCT	Chronic low back pain due to osteoarthritis	Non-individualised homeopathy, oral	Placebo	Physiotherapy	Pain Mobility/ROM Disability Medication use
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found.						
Homeopathy versus 'other' intervention**						
Stam 2001 (160)	RCT	Low back pain, acute	Non-individualised homeopathy, topical	Active control	Paracetamol	Pain Medication use

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D8.2.2 Risk of bias per item

The risk of bias for each item in the included studies for back and neck pain is described below and shown graphically in Figure D-16 (details are provided in Appendix E).

Bias arising from the randomisation process

Two studies (Gupta 2020, Stam 2001) were assessed as low risk of bias for this domain. Both studies detailed the randomisation process, allocation sequence and concealment, and baseline characteristics were balanced between intervention and control groups.

There were some concerns of bias in one study (Morris 2016). The study detailed the randomisation process, allocation sequence and concealment, however, baseline pain medication use was not balanced between the intervention and control groups, suggesting possible issues with the randomisation process (although possible it was by chance).

Bias due to deviations from intended interventions

One study (Morris 2016) was assessed as low risk of bias for this domain as an appropriate analysis was used. One study (Gupta 2020) was assessed to have some concerns for this domain as an inappropriate analysis was used. However, the impact of this on the outcome is not expected to bias the results. One study (Stam 2001) was assessed as high risk for this domain as the method of analysis was inappropriate and likely excluded eligible participants which could have biased the results.

Bias due to missing outcome data

Three studies (Gupta 2020, Morris 2016, Stam 2001) were assessed as low risk of bias for this domain as data were available for all, or nearly all participants, the analysis addressed missing data, or the proportions of and reasons for missing participants were similar across groups.

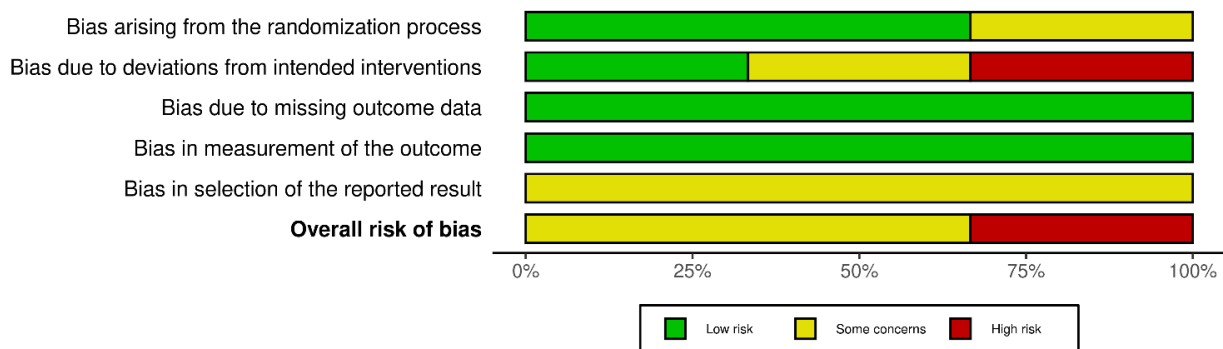
Bias in measurement of the outcome

Three studies (Gupta 2020, Morris 2016, Stam 2001) were assessed as low risk of bias for this domain as the methods of outcome assessment were appropriate and comparable between groups, and outcome assessors were unaware of the intervention received by study participants.

Bias in selection of the reported result

Three studies (Gupta 2020, Morris 2016, Stam 2001) were assessed to have some concerns for this domain as no pre-specified analysis plans were provided.

Figure D-17 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Back and/or neck pain



D8.2.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with back and neck pain are listed in Table D-34.

Table D-34 Outcomes considered by the NTWC to be critical or important for decision-making: Back and/or neck pain

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Gupta 2020	Morris 2016
Pain	VAS or McGill Pain Questionnaire	Critical	Yes	✓	✓ [^]
	CSPMS	Critical	Yes	✓	--
Physical function/mobility	ROM for lumbar spine	Critical	Yes	--	✓ [^]
	Stiffness VAS	Critical	Yes	✓	--
Disability	ODI or RMDS	Critical	Yes	--	✓ [^]
Quality of life	SF-36 or similar	Critical	Yes	✓	--
Medication use	Paracetamol use	Critical	Yes	--	✓ [^]

Abbreviations: CSPMS, cervical spondylosis pain management scale; ODI, Oswestry disability index; RMDS, Roland Morris disability scale; ROM, range of motion; SF-36, 36-item short form survey; VAS, visual analogue scale;

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results
- ? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study
- ^ Study data was not able to be extracted. Data were presented in figures and incomplete (missing values)

D8.2.3.1 Primary Comparison (vs placebo)

Two RCTs (Gupta 2020, Morris 2016) comparing homeopathy with placebo in people with back or neck pain were eligible for comparison. The studies contributed data all 5 outcomes considered critical for this review.

There were 4 studies awaiting classification (total 966 participants) that compared homeopathy with placebo that could have contributed to 4 of the 5 outcome domains considered critical for this review (see Appendix C4). Two studies were not available in English, one study was available as a conference abstract, and the results of one study were published after the search date. None were suspected of nonreporting of results for reasons related to the *p*-value, magnitude or direction of the results.

There were also 4 ongoing studies (not yet complete) comparing homeopathy with placebo (target 222 participants) that will likely contribute to at least 2 of the 5 outcome domains considered critical for this review (see Appendix C5).

Pain intensity

Two studies (Gupta 2020, Morris 2016) assessed pain intensity measured using a 10-cm VAS at the end of treatment (range 8 days to 6 weeks). The VAS is a subjective assessment of pain, reported by participants and measured on a continuous scale (mm) from 0 (no pain) to 100 (worst imaginable pain). Higher values indicate worse pain intensity. Among patients with subacute or chronic pain (including back or neck pain), the MCID is estimated to be 20 mm (154, 161, 162). For acute low back pain, the MCID is suggested to be 35 mm (161).

The results from one study (total 134 participants) suggested pain intensity was lower in the homeopathy group after a 7-day treatment period compared with the placebo group, but the difference between the groups was not significant (MD -0.74; 95% CI -1.88, 0.40; *p* = 0.20) (*GRADE: Low*).

Data were incomplete for one study (Morris 2016) and could not be included in the data synthesis. The authors noted there was improvement in median pain scores in both groups the end of treatment (6 weeks), but improvement was consistently better in the homeopathy group (*p* < 0.001).

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data.

Functional movement/Mobility

Two studies (Gupta 2020, Morris 2016) assessed mobility at the end of treatment (range 8 days to 6 weeks) using different measures.

One study (Gupta 2020) (total 134 participants) measured stiffness using a 10-cm VAS. The authors reported a significant improvement in stiffness in the homeopathy and placebo groups after a 7-day treatment period, but the difference between groups was not significant (MD -0.36; 95% CI -1.48, 0.76; *p* = 0.53) (*GRADE: Low*).

The study also assessed limitation of movement upon flexion, extension, side bending, and rotation as part of the Cervical Spondylosis Pain Management Scale (CSPMS). Information on the outcome measure and scoring were not available. The authors suggested participants in the homeopathy and placebo groups had a significant improvement in all outcomes after a 7-day treatment period, but the difference between groups was not significant as outlined below:

- flexion MD -0.10; 95% CI -0.32, 0.12; *p* = 0.38

- extension MD –0.08; 95% CI –0.29, 0.13; $p = 0.46$
- side bending MD –0.12; 95% CI –0.34, 0.10; $p = 0.29$
- rotation MD –0.02; 95% CI –0.27, 0.23; $p = 0.88$

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data.

Data were incomplete for one study (Morris 2016) (total 30 participants) and could not be included in the data synthesis. The authors measured range of motion (ROM) for flexion and extension using a tape measure, placed directly over the lumbar spine while the patient bends as far as they can (superior or inferior to a horizontal line). The authors reported improvement in the median ROM for flexion and extension in both groups at the end of treatment (6 weeks), but improvement was consistently better in the homeopathy group for both flexion ($p = 0.002$) and extension ($p = 0.021$).

Disability

One study (Morris 2016) measured disability using the Oswestry disability index (ODI) at the end of treatment (6 weeks). The ODI is used to quantify disability related to lower back pain. The questionnaire is comprised of 10 questions that assess the ability of people with low back pain to manage everyday life. Answers are scored on a 0 (no disability) to 5 (great deal of disability) scale. The final score ranges from 0 to 100, with a score of 0-20 indicating minimal disability, 21-40 indicates moderate disability, 41-60 indicates severe disability, 61-80 indicates crippled (back pain impinges on all aspects of life), and 81-100 indicating complete disability (bed-bound). In people with chronic low back pain the minimal important change is reported to be 12.88 (sensitivity 88%, specificity 85%) (163).

Data from Morris 2016 (total 30 participants) were incomplete and not able to be included in the data synthesis. The authors reported a reduction in the median ODI for the homeopathy and placebo groups after the 6-week treatment period, and suggested improvement was consistently better in the homeopathy group ($p < 0.001$). Median scores in both groups were below 20, suggestive of minimal disability. (GRADE: Very Low)

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data.

Quality of life

One study (Gupta 2022) measured quality of life using the Patient's Global Impression of Change Scale at the end of treatment (8 days). Participants were asked to rate quality of life on a 0–10 scale, where '0' was 'much better' and '10' was 'much worse'.

Results from one study (total 134 participants) suggested there was no difference in quality of life between the homeopathy and placebo groups (MD –0.64; 95% CI –1.35, 0.07; $p = 0.08$) (GRADE: Low).

Medication use

One study (Morris 2016) measured medication use, reported specifically as paracetamol use (tablets per week) at the end of treatment (6 weeks).

Data from Morris 2016 (total 30 participants) were incomplete and not able to be included in the data synthesis. The authors suggested there was no significant difference between treatment groups in the median quantity of pain medication used ($p = 0.531$). (GRADE: Very Low)

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data.

D8.2.3.2 Secondary Comparison (vs 'inactive' control)

There were no eligible studies found that compared homeopathy with inactive control in people with back and neck pain.

D8.2.3.3 Tertiary Comparison (vs other)

One RCT comparing homeopathy with 'other' interventions in people with back and neck pain was eligible for this comparison and contributed data for 2 of the 5 outcomes considered critical for this analysis. Data from this study is presented in Appendix F2 Supplementary outcome data.

There was one study awaiting classification (total 43 participants) that was published in language other than English (see Appendix C4).

D9 Diseases of the genitourinary system

D9.1 Menopausal symptoms or complaints

D9.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-35. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-35 Overview of PICO criteria of included studies: Menopausal symptoms or complaints

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Andrade 2019 (164)	RCT	Women with menopausal hot flushes	Non-individualised, oral (<i>Malagueta</i> [<i>Capsicum frutescens</i>])	Placebo	None reported	Symptom severity
Colau 2012 (165)	RCT	Women with menopausal hot flushes	Non-individualised, oral combination (Acthène)	Placebo	None reported	Symptom severity Hot flush severity HRQoL
Gupta 2019 (166)	RCT	Women with menopausal symptoms	Non-individualised, oral (Sepia 200C)	Placebo	None reported	Symptom severity
Jacobs 2005 (167)	RCT	Breast cancer survivors with menopausal symptoms	Individualised homeopathy, oral (single verum) OR Non-individualised homeopathy, oral combination (Hyland's menopause)	Placebo	All groups received consultations with a homeopath	Symptom severity Hot flush severity HRQoL
Von Hagens 2012 (168)	RCT, crossover	Women with menopausal symptoms	Non-individualised, oral combination (<i>Apis regina</i> GL D4, <i>Argentum metallicum</i> D5, <i>Ovaria bovis</i> GL D4)	Placebo	None reported	Symptom severity
Homeopathy versus inactive control (no intervention, waitlist or inactive usual care)*						
Relton 2012 (169)	RCT	Women with menopausal hot flushes	Individualised homeopathy, oral + attended up to 5 consultations with a homeopath	Inactive control (no intervention)	None reported	Symptom severity Hot flush severity HRQoL
Homeopathy versus 'other' intervention**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: HRQoL, health-related quality of life; RCT, randomised controlled trial.

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D9.1.2 Risk of bias per item

The risk of bias for each item in the included studies for menopausal symptoms is described below and shown graphically in Figure D-17 (details are provided in Appendix E).

Bias arising from the randomisation process

All 6 studies (Andrade 2019, Colau 2012, Gupta 2019, Jacobs 2005, Relton 2012, Von Hagens 2012) were assessed to have low risk of bias in this domain. These studies provided sufficient detail on the randomisation process, allocation concealment and baseline characteristics. There were no significant baseline differences noted in 2 studies (Colau 2012 and Gupta 2019), and in 4 studies (Andrade 2019, Jacobs 2005, Relton 2012 and Von Hagens 2012), some baseline imbalances were noted, but considered not likely due to the randomisation process.

Bias due to deviations from intended interventions

Six studies (Andrade 2019, Colau 2012, Gupta 2019, Jacobs 2005, Relton 2012, Von Hagens 2012) were assessed to have low concerns in this domain. Of these, 5 studies (Andrade 2019, Colau 2012, Gupta 2019, Jacobs 2005 and Von Hagens 2012) were placebo-controlled and reported no deviations from the intended intervention. One study (Relton 2012) reported deviations from the intended intervention that included refusal of treatment, probably due to the non-blinded nature of the study. These deviations were not considered due to the trial context. All studies performed ITT analysis, with some studies performing both ITT and per-protocol analysis.

Bias due to missing outcome data

One study (Gupta 2019) was assessed to have low concerns in this domain as outcome data was available for all randomised participants. Two studies (Colau 2012 and Relton 2012) were assessed as having some concerns in this domain due to the proportion of missing outcome data. Reasons for participant drop-out in these studies were not provided.

Three studies (Andrade 2019, Jacobs 2005 and Von Hagens 2012) were deemed to have high concerns in this domain. This was due to the proportion of missing outcome data and reasons for participant drop-out relating to their health status.

Bias in measurement of the outcome

Five studies (Andrade 2019, Colau 2012, Gupta 2019, Jacobs 2005, Von Hagens 2012) were assessed to have low concerns in this domain. These studies used valid outcome measures that did not differ between treatment groups.

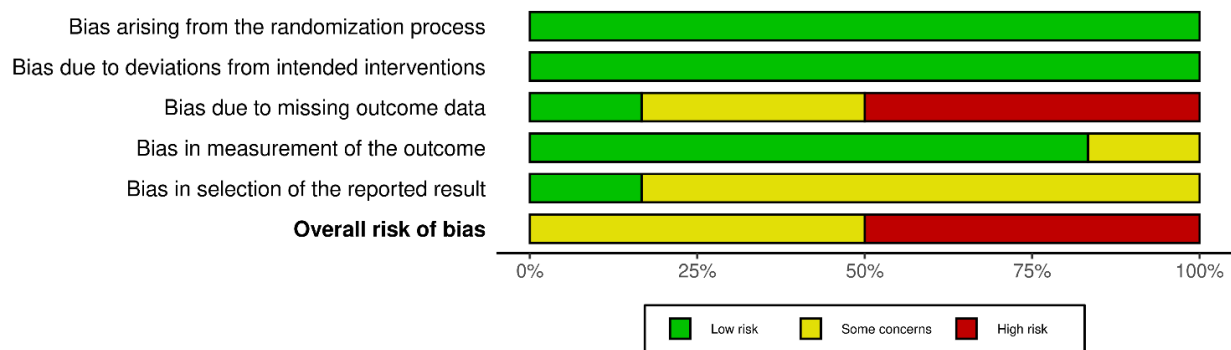
One study (Relton 2012) was assessed to have some concerns in this domain. In this study, the participants were aware of their treatment allocation, and the outcome measures were self-reported. It is possible that knowledge of the intervention could have influenced the assessment of the outcome, however, there was no evidence to suggest this occurred.

Bias in selection of the reported result

One study (Von Hagens 2012) was assessed to have low concerns in this domain as there was no evidence of multiple eligible outcome measurements or analysis, and a trial protocol was available for comparison.

Five studies (Andrade 2019, Colau 2012, Gupta 2019, Jacobs 2005, Relton 2012) were assessed to have some concerns in this domain, as their pre-specified analysis plans were not available. Some evidence of selective reporting was noted in Relton 2012, as results for some outcome measures were not presented.

Figure D-18 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Menopausal symptoms or complaints



D9.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with menopausal symptoms are listed in Table D-36.

Table D-36 Outcomes considered by the NTWC to be critical or important for decision-making: Menopausal symptoms or complaints

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Andrade 2019	Colau 2012	Gupta 2019	Jacobs 2005	Relton 2012	Von Hagens 2012
Symptom severity	MRS II, GCS, KMI	Critical	Yes	--	✓	✓	✓^	✓	✓
Hot flush severity	Hot flush severity score	Critical	Yes	✓^	✓	--	✓^	✓	--
Health-related quality of life	SF-36, HFRDIS	Critical	Yes	--	✓	--	✓^	--	--
Hot flush frequency	Number of hot flushes (diary)	Critical	No	--	--	--	--	--	--
Night sweat frequency	Daily diary	Critical	No	--	--	--	--	--	--

Abbreviations: GCS, Greene climacteric scale; HFRDIS, Hot Flash Related Daily Interference Scale; KMI, Kupperman menopausal index; MRS II, menopause rating scale; SF-36, short form-36.

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

--No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

^ Results not in extractable form

D9.1.3.1 Primary Comparison (vs placebo)

Five RCTs (Andrade 2019, Colau 2012, Gupta 2019, Jacobs 2005, Von Hagens 2012) comparing homeopathy with placebo in people with menopausal symptoms or complaints were eligible for this comparison and contributed data to 3 outcomes considered critical or important for this review.

There was one study awaiting classification (35 participants) that compared homeopathy with placebo that could have contributed data to one critical outcome domain, but full study results are not published. There were also 2 ongoing studies not yet complete comparing homeopathy with placebo (target 104 participants) that could contribute to 2 outcome domains critical for this review (see Appendix C5)

Symptom severity

Four studies (Colau 2012, Gupta 2019, Jacobs 2005, Von Hagens 2012) measured symptom severity using a range of outcome measures at the end of treatment (range 12 weeks to 12 months).

Two studies (Colau 2012, Von Hagens 2012) used the Menopause Rating Scale (MRS): a self-administered tool consisting of 11 items relating to 3 domains (somatic, psychological, urogenital). Each item is rated on a 5-point Likert scale, ranging from 0 (no complaints) to 4 (severe symptoms). With the total score being a sum of all items (maximum 44 points). The MCID for MRS is not established. Both studies (total 204 participants) suggested improvements in the MRS were observed in both the homeopathy and placebo groups, but the difference between groups was not significant (SMD -0.26; 95% CI -0.03, 0.54; $p = 0.07$; $I^2 = 0\%$).

One study (Gupta 2019) measured symptom severity using the Greene Climacteric Scale (GCS) at the end of treatment (6 months). The GSC measures 21 symptoms of menopause on a scale from 0 (none) to 3 (severe). Three main areas are measured including psychological (items 1 to 11), physical (items 12 to 18) and vasomotor (items 19 and 20), with item 21 related to sexual function. The total score comprises the sum of severity scores (maximum score 63). The MCID for GCS scores could not be identified, but a score over 12 on the GCS suggests menopausal symptoms that interfere with daily living. Results from this study (88 participants) suggested a reduction in symptom severity favouring homeopathy compared with placebo (SMD -0.72; 95% CI -1.15, -0.29; $p = 0.001$).

One study (Jacobs 2005) measured symptom severity using the Kupperman Menopausal Index (KMI). Results from this study were not able to be included in the meta-analysis, as the study authors did not report any usable data. Study authors reported a positive trend towards a lower KMI score in the individualised homeopathy group compared to placebo ($p = 0.10$), but also note an increase in headaches in the non-individualised homeopathy group ($p = 0.04$).

Pooled results from 3 studies (total 290 participants) with available data suggested there is no difference in symptom severity comparing homeopathy with placebo (SMD -0.07; 95% CI -0.70, 0.56; $p = 0.83$; $I^2 = 86\%$). (GRADE: Very Low)

In a sensitivity analysis conducted to assess the impact of one study at high risk of bias, the confidence in the results did not change (SMD -0.19; 95% CI -1.22, 0.85; $p = 0.73$; $I^2 = 92\%$).

Hot flush severity

Three studies (Andrade 2019, Colau 2012, Jacobs 2005) reported hot flush severity measured using 2 different measures at the end of treatment (range: 4 weeks to 12 months).

One study (Andrade 2019) (total 40 participants) measured hot flush severity using the Measure Yourself Medical Outcome Profile (MYMOP), with hot flushes the designated symptom to be assessed. Participants rate symptoms on scale from 0 ("as good as it can be") to 6 ("as bad as it could be"). The study authors did not provide baseline or end of treatment scores, but reported there was no difference between the homeopathy and placebo groups (MD -0.06; 95% CI -0.66, 1.86; $p = 0.07$)

Two studies (Colau 2012, Jacobs 2005) used a hot flush severity score (HFS), which is defined as the daily frequency of hot flushes multiplied by flush intensity (graded from 1 [mild] to 4 [very strong]). The measure is self-reported, using a daily diary. The MCID for HFS in people with menopausal symptoms is estimated to be a weekly reduction of 25 hot flushes (170). Results from one study (Colau 2012) (108 participants) suggested the reduction hot flush score at the end of study (12 weeks) was significantly higher in the homeopathy group compared with placebo (MD -30.70; 95% CI -57.66, -3.74; $p = 0.03$).

(GRADE: Very Low)

Results from Jacobs 2005 (total 83 participants) were not able to be included in the meta-analysis as the data were incomplete. The study authors reported no significant difference in hot flush severity in an adjusted univariate model comparing individualised (single) homeopathy with placebo (MD -12.0; 95% CI -34.3, 10.3; $p = 0.3$) or comparing non-individualised (combination) homeopathy with placebo (MD -0.4; 95% CI -22.3, 10.3; $p = 1.0$).

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data.

Health-related quality of life

Two studies (Colau 2012, Jacobs 2005) measured health-related quality of life using the hot flush related daily interference scale (HFRDIS) or the SF-36 scale at the end of treatment (range 12 weeks to 12 months). Results from Jacobs 2005 were not able to be included in the meta-analysis as the authors did not report complete data⁷.

The HFRDIS is a 10-item scale measuring the degree hot flashes interfere with 9 daily activities, with the last item assessing overall quality of life. Participants rate the degree to which hot flashes have interfered with each item during the previous week from 0 (do not interfere) to 10 (completely interfere). The MCID for HFRDIS is estimated to be a reduction of 1.66 points (171).

Results from one study (Colau 2012) (108 participants) suggested no important difference in HFRDIS scores between the homeopathy and placebo groups at the end of treatment (12 weeks) (MD -0.30; 95% CI -0.65, 1.25; $p = 0.53$) (GRADE: Low).

No sensitivity analysis was conducted to assess the impact of studies at high risk of bias, as only one study contributed data.

The SF-36 is a self-reported questionnaire consisting of 8 domains that are summarised on a scale from 0 (worse) to 100 (best). The SF-36 can be summarised into 2 component scores. The physical component summary (PCS) score includes the domains of general health, physical functioning, role physical and body pain. The mental component summary (MCS) score includes the domains of vitality, social functioning, role emotional, and mental health. The PCS and MCS are derived by aggregating individual scores. The MCID for the SF-36 is estimated to be around 2 to 4 points for the general population (i.e. ~0.5 of the SD) (172).

D9.1.3.2 Secondary Comparison (vs 'inactive' control)

One RCT (Relton 2012) (43 participants) comparing homeopathy with an inactive control (no intervention) in people with menopausal symptoms was eligible for this comparison. The study contributed to 2 outcome domains considered critical for this review.

There were no ongoing studies and no studies awaiting classification that compared homeopathy with inactive control (no intervention).

⁷ Authors selectively reported p-values for outcome domains suggestive of significant effect favouring homeopathy.

Symptom severity

One study (Relton 2012) measured symptom severity using the Greene Climacteric Scale (GCS) and the Measure Yourself Medical Outcome Profile (MYMOP) at the end of treatment (36 weeks). The GCS results were included here as it is the preferred (multidimensional) outcome measure.

The GSC measures 21 symptoms of menopause on a scale from 0 (none) to 3 (severe). Three main areas are measured including psychological (items 1 to 11), physical (items 12 to 18) and vasomotor (items 19 and 20), with item 21 related to sexual function. The total score comprises the sum of severity scores (maximum score 63). The MCID for GCS scores could not be identified, but a score over 12 on the GCS suggests menopausal symptoms that interfere with daily living.

The results from one study (43 participants) suggested there was no significant difference in the mean change from baseline for symptom severity between the homeopathy and control (no intervention) groups (MD -3.78; 95% CI -7.84, 0.28; $p = 0.07$). (*GRADE: Very Low*).

Hot flush severity

One study (Relton 2012) measured hot flush severity using the hot flush frequency and severity score (HFS), which is defined as the daily frequency of hot flushes multiplied by flush intensity (graded from 1 [mild] to 4 [very strong]) (173). The measure is self-reported, using a daily diary. The MCID for HFS in people with menopausal symptoms is unknown, with a weekly reduction of 25 or more moderate to severe hot flushes considered clinically important (170).

Results from one study (43 participants) suggested a greater reduction in HFS in the homeopathy group compared with the control (no intervention) group (MD -5.73; 95% CI -11.94, 0.48; $p = 0.07$). (*GRADE: Very Low*).

D9.1.3.3 Tertiary Comparison (vs other)

No studies comparing homeopathy with 'other' interventions in people with menstrual disorders were identified.

D9.2 Menstrual disorders

D9.2.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-37. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-37 Overview of PICO criteria of included studies: Menstrual disorders

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Charandabi 2016 (174, 175)	RCT	Dysmenorrhea	Individualised homeopathy	Placebo	Both groups received consultations with a homeopath	Pain Medication use HRQoL
Singh 2020 (176)	RCT	Dysmenorrhea	Individualised homeopathy	Placebo	Both groups received consultation with a homeopath	Pain
Teixeira 2016 (177)	RCT	Endometriosis	Non-individualised homeopathy, oral (potentised estrogen 12cH, 18cH or 24cH)	Placebo	Both groups received consultations with a homeopath	Pain HRQoL Anxiety Depression
Yakir 1994 (178, 179)	RCT	Premenstrual syndrome	Individualised homeopathy, oral	Placebo	Both groups received consultations with a homeopath	Symptom severity Medication use
Yakir 2019 (180)	RCT	Premenstrual syndrome	Individualised homeopathy	Placebo	Both groups received consultations with a homeopath	Symptom severity Medication use
Homeopathy versus inactive control (no intervention, waitlist or inactive usual care)						
Klein-Laansma 2017 (181)	RCT	Premenstrual syndrome and premenstrual disorder	Individualised homeopathy	Inactive control (no intervention)	Usual care	Pain
Homeopathy versus 'other' intervention**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: HRQoL, health-related quality of life; RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D9.2.2 Risk of bias per item

The risk of bias for each item in the included studies for menstrual disorders is described below and shown graphically in Figure D-18 (details are provided in Appendix E).

Bias arising from the randomisation process

All 6 studies (Charandabi 2016, Klein-Laansma 2017, Singh 2020, Teixeira 2016, Yakir 2019 and Yakir 1994) were assessed to have low concerns in this domain. Three studies (Charandabi 2016, Klein-Laansma 2017 and Teixeira 2017) had some slight baseline imbalances, but were deemed not likely due to the randomisation process. Two studies (Singh 2020 and Yakir 1994) did not present baseline characteristics, however it was presumed they were balanced due to their sound randomisation processes. One study (Yakir 2019) reported no significant baseline differences between the 2 treatment arms.

No issues were identified in any of the studies relating to allocation sequence and concealment.

Bias due to deviations from intended interventions

All 6 studies (Charandabi 2016, Klein-Laansma 2017, Singh 2020, Teixeira 2016, Yakir 2019 and Yakir 1994) were assessed to have low concerns in this domain. Five studies (Charandabi 2016, Singh 2020, Teixeira 2016, Yakir 2019 and Yakir 1994) were double-blind trials with no reported deviations due to the trial context. One study (Singh 2020) conducted an ITT analysis, and 2 studies (Charandabi 2016 and Yakir 1994) used a modified-ITT analysis. Two studies (Teixeira 2016 and Yakir 2019) conducted both ITT and per-protocol methods of analysis.

One study (Klein-Laansma 2017) was a single-blinded study, where only the researchers were blinded to the treatment allocations. In this study, both ITT and per-protocol analysis was performed.

Bias due to missing outcome data

One study (Teixeira 2016) was deemed to have high concerns in this domain due to the proportion of missing outcome data and reasons for participant drop-out relating to health status.

Four studies (Charandabi 2016, Klein-Laansma 2017, Yakir 2019 and Yakir 1994) were assessed as having some concerns in this domain due to missing outcome data. In these studies, health status was not provided as a reason for participant drop-out. In Yakir 1994, participant drop-out numbers were consistent across the treatment arms. In Charandabi 2016, Klein-Laansma 2017 and Yakir 2019, participant drop-out was not evenly distributed across treatment arms.

One study (Singh 2020) was assessed to have low concerns in this domain as outcome data was provided for almost all (98%) of the randomised participants.

Bias in measurement of the outcome

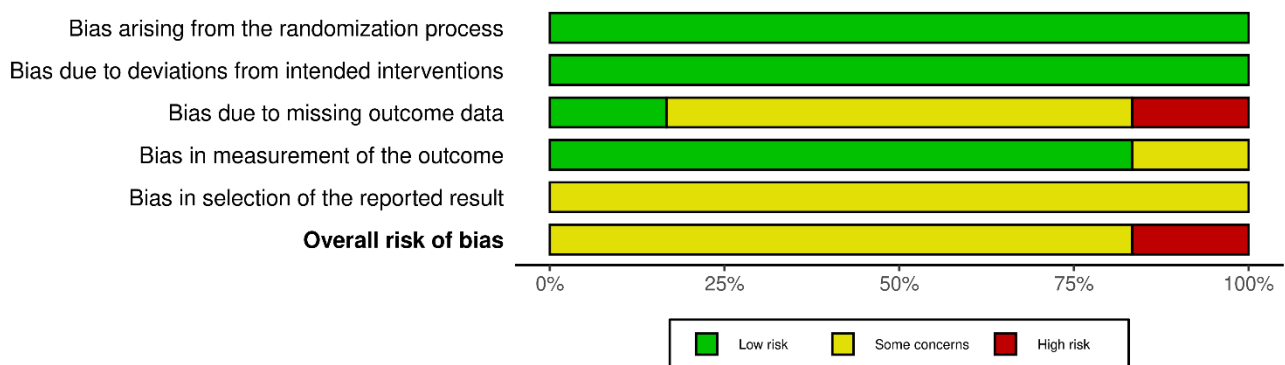
Five studies (Charandabi 2016, Singh 2020, Teixeira 2016, Yakir 2019 and Yakir 1994) were deemed to have low concerns in this domain. These studies used valid outcome measures that did not differ between treatment arms, and the outcome assessors were blinded to treatment allocation.

One study (Klein-Laansma 2017) was assessed as having some concerns in this domain. In this study, the participants were aware of their treatment allocation, and the outcome measures were self-reported. It is possible that knowledge of the intervention could have influenced the assessment of the outcome, however, there was no evidence to suggest this had occurred.

Bias in selection of the reported result

All 6 studies (Charandabi 2016, Klein-Laansma 2017, Singh 2020, Teixeira 2016, Yakir 2019 and Yakir 1994) were assessed to have some concerns in this domain. Five studies (Charandabi 2016, Klein-Laansma 2017, Singh 2020, Yakir 2019 and Yakir 1994) did not have a pre-specified analysis plan available for comparison. One study (Teixeira 2016) had a pre-specified analysis plan, however there was some evidence of selective reporting of outcomes, as only 3 of 8 quality-of-life outcome domains were assessed.

Figure D-19 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Menstrual disorders



D9.2.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with menstrual disorders are listed in Table D-38.

Table D-38 Outcomes considered by the NTWC to be critical or important for decision-making: Menstrual disorders

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Charandabi 2016	Klein-Laansma 2017	Singh 2020	Teixeira 2016	Yakir 2019	Yakir 1994
Pain intensity	VAS	Critical	Yes	✓	--	✓	✓^	--	--
Symptom severity	MDQ	Critical	Yes	--	✓	--	✓^	✓	✓
Depression	BDI	Critical	Yes	--	--	--	✓^	--	--
Anxiety	BAI	Critical	Yes	--	--	--	✓^	--	--
Health-related quality of life	SF-36	Critical	Yes	✓	--	--	✓^	--	--
Medication use	Number of medications	Critical	Yes	✓	--	--	--	✓	✓^

Abbreviations: BAI, beck anxiety inventory; BDI, beck depression inventory; MDQ, menstrual distress questionnaire; SF-36, 36-item short form; VAS, visual analogue scale.

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the p-value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

^ Study data were not able to be extracted.

D9.2.3.1 Primary Comparison (vs placebo)

Five RCTs (Charandabi 2016, Singh 2020, Teixeira 2016, Yakir 1994, Yakir 2019) comparing homeopathy with placebo in people with menstrual disorders were eligible for this comparison. The studies contributed data to all 6 outcomes considered critical or important for this review.

There were 4 studies awaiting classification (total 252+ participants) and 8 ongoing studies (target 615 participants) that compared homeopathy with placebo that could have contributed data to at least 4 critical outcome domains. Complete results from 5 studies (292+ participants) were judged likely to be missing because the *p*-value, magnitude or direction of the results generated were considered unfavourable by the study investigators (see Appendix C6).

No sensitivity analysis was conducted that examined the impact of RCTs at high risk of bias, as only one study was at high risk of bias, and it did not contribute any data to the analysis.

Pain

Three studies (total 162 participants) reported pain using a Visual Analog Scale (VAS) at the end of treatment (range: 3 to 6 months) (Charandabi 2016, Singh 2020, Teixeira 2016). Results from one study (Teixeira 2016) were not able to be included in the meta-analysis, as the study did not report data for the placebo group.

The VAS is a validated tool used for the subjective measurement of pain on a continuous scale from 0 to 10 (or 0 to 100), with higher scores indicating worse pain. The MCID in VAS pain scores for people with menstrual disorders was not identified, but the median absolute MCID in people with chronic pain is reported to be 20 mm (IQR 15–30) (154).

Pooled results from 2 studies (total 112 participants) suggested there is no important difference in mean pain scores between the homeopathy and placebos (MD –15.25; 95% CI –36.49, 5.98; *p* = 0.16; *I*² = 88%) (*GRADE: Very Low*).

Symptom severity

Three studies (total 165 participants) reported symptom severity using a modified Moos Menstrual Distress Questionnaire (MDQ) or the endometriosis-associated pelvic pain (EAPP) at the end of treatment (3 months) (Teixeira 2016, Yakir 1994, Yakir 2019). Results from one study (Teixeira 2016) were not able to be included in the meta-analysis, as the study did not report data for the placebo group.

The MDQ is a self-reporting tool, used by participants to measure the severity of their pre-menstrual symptoms. It is comprised of 37 or 38 predefined symptoms in 6 categories that included pain, functioning, appetite, autonomic reactions, water retention and mental symptoms. Symptom severity scores are rated daily on a scale from 0 to 4, with higher scores indicating greater symptom severity. The MCID for the MDQ for people with menstrual disorders could not be identified.

Pooled results from 2 studies (total 115 participants) suggested the improvements in symptom severity were not different between the homeopathy and placebo groups (MD –0.10; 95% CI –0.25, 0.04; *p* = 0.16; *I*² = 34%) (*GRADE: Very Low*).

Depression

One RCT (Teixeira 2016) (41 participants analysed) reported depression measured using the Beck Depression Inventory (BDI) at the end of treatment (3 months). The results were not able to be included in the meta-analysis, as the data were incomplete.

The BDI assesses the behavioural and cognitive symptoms of depression and consists of 21 questions, each on a 4-point scale. Scores range from 0 to 63 with a higher score indicating a greater level of depressive symptoms.

The study authors suggested significant improvement from baseline in the homeopathy group (*p* < 0.001). No data for the placebo group were reported, with the authors noting the BDI was not balanced at baseline, being significantly higher in the placebo group (*p* = 0.004).

Anxiety

One RCT (Teixeira 2016) (41 participants analysed) measured anxiety using the Beck Anxiety Inventory (BAI) at the end of treatment (3 months). The results were not able to be included in the meta-analysis, as the data were incomplete.

The BAI is a 21 item self-reported inventory measuring the severity of anxiety symptoms in adults. Each item is scored on a scale from 0 (not at all) to 3 (severely) to yield a total score from 0 to 63 where a higher score indicates more severe anxiety.

The study authors suggested significant improvement from baseline in anxiety scores for both the homeopathy and placebo groups ($p = 0.001$).

Health-related quality of life

Two RCTs (Charandabi 2016, Teixeira 2016) reported health-related quality of life measured using the short-form (SF-36) health survey. Results from one study (Teixeira 2016) were not able to be included in the meta-analysis, as the data were incomplete and study did not report data for the placebo group.

The SF-36 is a questionnaire consisting of 8 domains that are summarised on a scale from 0 (worse) to 100 (best). The SF-36 can be summarised into 2 component scores. The physical component summary (PCS) score includes the domains of general health, physical functioning, role physical and body pain. The mental component summary (MCS) score includes the domains of vitality, social functioning, role emotional, and mental health. The PCS and MCS are derived by aggregating individual scores. The MCID for the SF-36 is estimated to be around 2 to 4 points for the general population (i.e. ~ 0.5 of the SD) (172).

Results from one study (Charandabi 2016) (total 47 participants) suggested there is no significant difference in quality of life between the homeopathy and placebo groups for both the PCS (MD -0.50 ; 95% CI $-8.50, 7.50$; $p = 0.90$) or the MCS (MD 4.60 ; 95% CI $-3.50, 12.70$; $p = 0.27$). (GRADE: Very Low)

Medication use

Three RCTs (Charandabi 2016, Yakir 1994, Yakir 2019) reported medication use, measured using different criteria.

One study (Charandabi 2016) (total 54 participants) reported the number of nonsteroidal anti-inflammatory (NSAID) pills consumed by participants during their menstrual cycle. Results after one treatment cycle suggested there was no difference in NSAID use between the homeopathy and placebo groups (SMD 0.00 ; 95% CI $-0.58, 0.58$).

One study (Yakir 1994) (total 23 participants) reported the number of medications consumed during the 7-day period prior to menstruation, however the results were not able to be included in the meta-analysis as data were incomplete. The study authors reported a reduction in medication use during the 3-months post treatment in both groups, with the difference between groups not significant (p -value not reported).

One study (Yakir 2019) (total 105 participants) reported the number of additional medications consumed during the 12 days before menstruation prior to treatment, and over 3 months post-treatment. Results for medication use post-treatment suggested there was no difference between the homeopathy and placebo groups (SMD -0.36 ; 95% CI $-0.76, 0.05$; $p = 0.06$).

Pooled results from 2 studies (143 participants) suggested there is no important difference between treatment groups (SMD -0.24 ; 95% CI $-0.57, 0.09$; $p = 0.16$; $I^2 = 0\%$). (GRADE: Very Low)

D9.2.3.2 Secondary Comparison (vs 'inactive' control)

One RCT (Klein-Laansma 2017) comparing homeopathy with inactive control (no intervention) in people with menstrual disorders was eligible for this comparison. The study contributed data to one outcome considered critical or important for this review.

There was one study awaiting classification (80 participants) that compared homeopathy with inactive control (no intervention) that could have contributed to 2 critical outcome domains, but the publication was not able to be retrieved. There were no ongoing studies that compared homeopathy with inactive control in people with menstrual disorders.

Symptom severity

One RCT (Klein-Laansma 2017) (60 participants) measured symptoms severity using the premenstrual tension syndrome visual analogue scale (PMTS-VAS) and the daily record of severity of problems (DRSP) at the end of treatment (4 months).

The PMTS-VAS measures 12 premenstrual symptoms on a scale from 0-100, with higher scores indicating greater symptom severity. Results for this study were not able to be included in the meta-analysis as the study did not report any data using this outcome measure.

The DRSP records 21 premenstrual symptoms as well as the impact on daily activities, for a total of 24 items. Participants are required to prospectively monitor and record symptom severity/impact for two consecutive menstrual cycles, from 1 (not at all) to 6 (extreme). Total scores range from 168 to 1008, with higher scores indicating greater impact.

The results from one study suggested a greater reduction in symptom severity in the homeopathy group when compared with no intervention (MD -125.00; 95% CI -198.26, -51.74; $p = 0.0008$). (GRADE: Low)

D9.2.3.3 Tertiary Comparison (vs other)

There were no studies comparing homeopathy with 'other' interventions in people with menstrual disorders that were eligible for this comparison.

D10 Symptoms, signs or clinical findings, not elsewhere classified

D10.1 Fatigue conditions

D10.1.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-39. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-39 Overview of PICO criteria of included studies: Fatigue conditions

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
McKendrick 1999 (182-185)	RCT	Chronic fatigue syndrome	Individualised homeopathy	Placebo	1 x consultation per month for 6 months	Fatigue Fatigue impact Health-related quality of life
Homeopathy versus inactive control (no intervention, waitlist, inactive usual care)*						
No studies found						
Homeopathy versus 'other' intervention**						
No studies found						

Note: CH=centesimal dilutions using Hahnemann's dilution method; D=Decimal dilutions; M=millesimal dilutions; X=Decimal dilutions

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D10.1.2 Risk of bias per item

The risk of bias for each item in the included studies for fatigue conditions is described below and shown graphically in Figure D-19 (details are provided in Appendix E).

Bias arising from the randomisation process

McKendrick 1999 was assessed at low risk of bias due to the randomisation process. Randomisation was by computer generated random numbers and concealed from the trialists. Baseline characteristics between groups suggested no issue with randomisation.

Bias due to deviations from intended interventions

McKendrick 1999 was judged at low risk of bias due to deviations from intended interventions, as participants and study staff were blind to treatment allocation, and an appropriate method of analysis was used.

Bias due to missing outcome data

McKendrick 1999 was assessed to have some concerns due to the proportion of missing outcome data. Data was missing for 16.5% of randomised participants, and no analysis to adjust for missing data was presented. The proportion of participants missing data was balanced between groups,

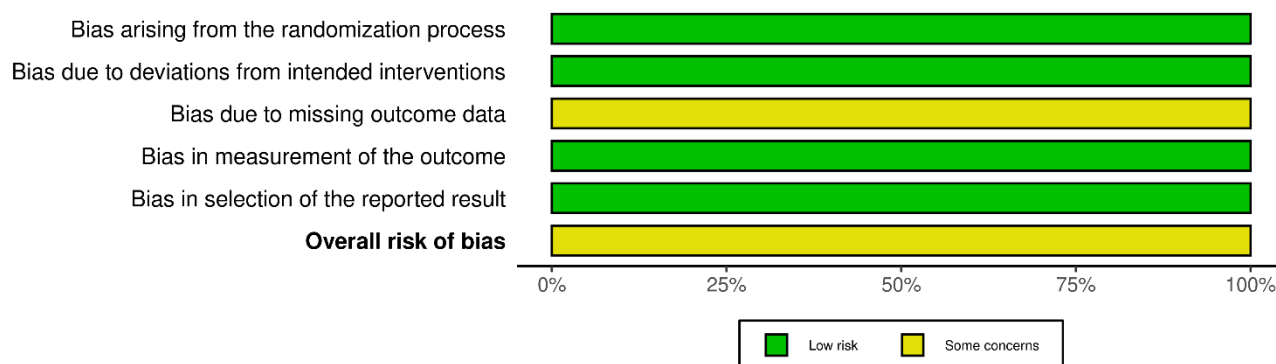
Bias in measurement of the outcome

McKendrick was assessed at low risk of bias for this domain, as appropriate measurement tools were used to assess the outcome, and outcome assessors were blinded to intervention group.

Bias in selection of the reported result

McKendrick was assessed at low risk of bias of the reported result, as analysis was conducted by a blinded statistician and the study was registered. There was no evidence of inappropriate selection or analysis of results.

Figure D-20 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – fatigue conditions



D10.1.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with fatigue conditions are listed in Table D-40.

Table D-40 Outcomes considered by the NTWC to be critical or important for decision-making: Fatigue conditions

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	McKendrick 1999
Fatigue	Multidimensional Fatigue Inventory (or other validated measure)	Critical	Yes	✓
Health-related quality of life	SF-36 (or other validated measure)	Critical	Yes	✓

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

D10.1.3.1 Primary Comparison (vs placebo)

One RCT (McKendrick 1999) comparing homeopathy with placebo in people with chronic fatigue syndrome was eligible for this comparison. The RCT contributed data to 2 outcomes considered critical or important for this review.

There were 3 ongoing studies comparing individualised homeopathy with placebo in people with post-COVID-19 fatigue (total 197 participants) that could have contributed data to 2 outcomes. Results of these studies are not yet published (see Appendix C5).

Fatigue

One RCT (86 participants) reported fatigue measured using the Multidimensional Fatigue Inventory (MFI) as a change from baseline to end of treatment (6 months) (McKendrick 1999).

The MFI is a 20-item self-report tool designed to measure fatigue over 5 domains: General Fatigue, Physical Fatigue, Mental Fatigue, Reduced Activity, and Reduced Motivation (186). Total scores range from 4 to 20, with a higher score indicating worse fatigue. No MCID in people with chronic fatigue was identified.

Result showed no significant difference in the change score for any of the 5 domains of the MFI as follows: (*GRADE: Low*)

- general fatigue (MD -1.35; 95% CI -2.77, 0.07; $p = 0.06$)
- physical fatigue (MD -0.85; 95% CI -2.30, 0.60; $p = 0.25$)
- mental fatigue (MD -0.65; 95% CI -2.12, 0.82; $p = 0.39$)
- reduced activity (MD -0.91; 95% CI -2.49, 0.67; $p = 0.26$)
- reduced motivation (MD 0.30; 95% CI -1.23, 1.83; $p = 0.70$)

No sensitivity analysis was conducted that examined the impact of RCTs at high risk of bias, as only one study was included in the analysis.

Quality of life

One RCT (86 participants) reported health-related quality of life measured using the Functional Limitations Profile (FLP) as a change from baseline to end of treatment (6 months) (McKendrick 1999).

The FLP is a health-related quality of life measure adapted from the Sickness Impact Profile. It contains 136-items and assesses difficulties in physical, psychological and social functioning across 12 domains. The physical dimension score is calculated as a percentage of the maximum possible dysfunction score across 4 categories: ambulation, body care and movement, mobility, and household management. The psychosocial domain is calculated as a percentage of the maximum possible dysfunction score across 5 categories: recreation and pastime, social interaction, emotion, alertness, and sleep and rest. Scores for the eating, communication, and work categories are reported separately. No MCID in people with chronic fatigue was identified.

Result showed no significant difference in the change score for either the physical or psychosocial domains as follows (*GRADE: Low*):

- physical (MD -2.39; 95% CI -6.03, 1.25; $p = 0.20$)
- psychosocial (MD -3.05; 95% CI -8.36, 2.26; $p = 0.26$)

No sensitivity analysis was conducted that examined the impact of RCTs at high risk of bias, as only one study was included in the analysis.

D10.1.3.2 Secondary Comparison (vs 'inactive' control)

There were no studies identified which compared homeopathy with inactive control in people with fatigue conditions.

D10.1.3.3 Tertiary Comparison (vs other)

There were no studies identified which compared homeopathy with other (active) interventions in people with fatigue conditions.

D10.2 Fibromyalgia

D10.2.1 List of studies

An overview of the PICO criteria of included studies is provided in Table D-1. Study details, including all outcome domains and measures reported by the included studies are provided in [Appendix F1](#). Outcome data for critical or important outcomes are provided in [Appendix F2](#).

Table D-41 Overview of PICO criteria of included studies: Fibromyalgia

STUDY ID	Study design	POPULATION	INTERVENTION	COMPARATOR	CO-INTERVENTION	OUTCOME DOMAINS
Homeopathy versus placebo*						
Bell 2004 (187-190)	RCT	Fibromyalgia	Individualised homeopathy, oral + Visits with a homeopath every 2 months for 6 months with optional crossover at 4 months	Placebo	Both groups received consultation with homeopath	Pain HRQoL Fatigue Emotional function Functional wellbeing
Fisher 1988 (191, 192)	Quasi RCT	Fibromyalgia	Non-individualised homeopathy, oral	Placebo	None	Pain Sleep HRQoL
Homeopathy versus inactive control (no intervention, waitlist or inactive usual care)*						
Relton 2009 (193)	RCT	Fibromyalgia	Individualised homeopathy, oral, x1 initial consultation with a homeopath followed by up to 4 interviews (4-6 weeks apart)	Inactive control (no intervention)	Usual care	Pain Fatigue HRQoL Emotional function
Homeopathy versus 'other' intervention**						
No studies found						

Abbreviations: RCT, randomised controlled trial

* Studies that compared homeopathy with placebo or an inactive control were eligible for inclusion in the evidence synthesis and are included in the Summary of findings tables if they reported outcomes considered critical or important to this review.

** Studies that compared homeopathy with an active intervention are included in the supplementary outcome tables ([Appendix F2](#)) if they reported data for outcomes considered critical or important to this review.

D10.2.2 Risk of bias per item

The risk of bias for each item in the included studies for fibromyalgia is described below and shown graphically in Figure D-20 (details are provided in Appendix E).

Bias arising from the randomisation process

Two studies (Bell 2004, Relton 2009) were assessed to have low risk of bias for this domain as sufficient detail regarding the randomisation sequence, allocation concealment and baseline characteristics were provided. Some baseline imbalances were noted in Bell 2004; however, they were considered unlikely due to the randomisation process. One study (Fisher 1988) was assessed as having some concerns in this domain, as details relating to the randomisation sequence and baseline characteristics were not presented.

Bias due to deviations from intended interventions

Two studies (Bell 2004, Relton 2009) were assessed at low risk of bias for this domain. In Bell 2004, both participants and study staff were blinded to treatment allocation, and an appropriate method of analysis was used (modified ITT). In Relton 2009, participants and study staff were not blinded, but there were no deviations from the intended intervention that were considered to have arisen due to the trial context. ITT analysis was specified and performed.

One study (Fisher 1988) was assessed to have some concerns in this domain due to insufficient information provided regarding the method of analysis and potential for bias.

Bias due to missing outcome data

Two studies (Bell 2004, Relton 2009) were assessed as having some concerns in this domain. These concerns were due to the proportion of missing outcome data in both studies, with no analysis presented to account for missingness. One study (Fisher 1988) was deemed to have high concerns of bias due to lack of information on missing outcome data. In this study, it was not specified whether there was any missing data, and the number of participants analysed was not provided.

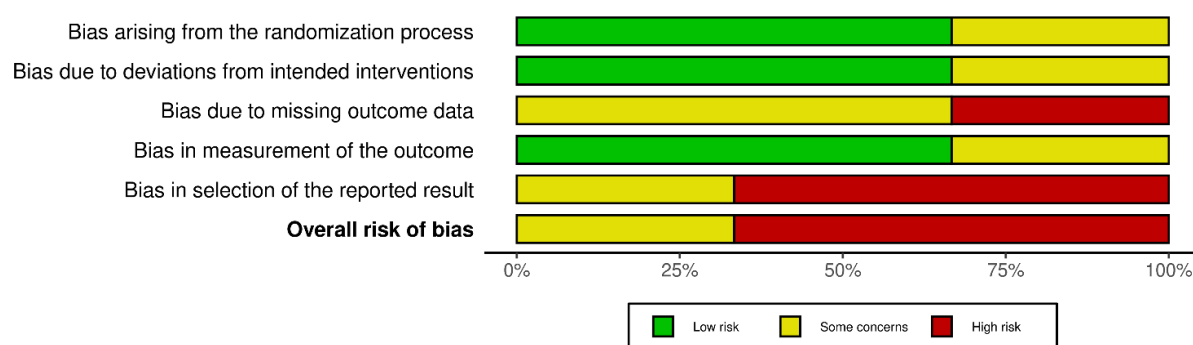
Bias in measurement of the outcome

Two studies (Bell 2004, Fisher 1988) were assessed as having low concerns in this domain. In these studies, the methods of outcome measurement were appropriate and did not differ between treatment groups. There were some concerns in one study (Relton 2009), as the participants were aware of their treatment allocation and were self-reporting most of their outcomes. It is possible that knowledge of the intervention could have biased self-reported outcomes, but there was no evidence to suggest that this is likely.

Bias in selection of the reported result

One study (Relton 2009) was assessed as having some concerns in this domain, as there was no pre-specified analysis plan. However, there was no evidence of selective reporting of outcomes based on multiple eligible outcomes or multiple analyses. Two studies (Bell 2004, Fisher 1988) were assessed to have high risk of bias. Bell 2004 had evidence of selective reporting of outcomes based on multiple eligible measures or domains, with not all domains of the POMS being reported and missing outcome data for the FACIT. Fisher 1988 was assessed to have high risk of bias, as continuous outcomes were transformed into binary variables for analysis, which created uncertainty as to whether the results were clinically meaningful or appropriate. It was unclear whether this method of data transformation was part of the study protocol as there was no pre-specified analysis plan.

Figure D-21 Risk of bias summary: review authors' judgements about each risk of bias item expressed as percentages across all RCTs – Fibromyalgia



D10.2.3 Effect of intervention

Outcomes considered by the NTWC to be critical or important for decision-making in people with fibromyalgia are listed in Table D-2.

Table D-42 Outcomes considered by the NTWC to be critical or important for decision-making: Fibromyalgia

Outcome domain	Measured with	Consensus rating	Data available for Primary or Secondary Comparison?	Bell 2004	Fisher 1988	Relton 2009
Pain intensity	McGill Pain Questionnaire or other validated measure	Critical	Yes	✓	✓ [^]	✓
Fatigue	FIQ – fatigue or other validated measure	Critical	Yes	✓	--	✓
Health-related quality of life	FIQ or other validated measure	Critical	Yes	✓	✓ [^]	✓
Emotional wellbeing	HADS, POMS or other validated measure	Critical	Yes	✓	--	✓
Pain disability	BPI – interference	Critical	No	--	--	--

Abbreviations: BPI, Brief Pain Inventory; FIQ, Fibromyalgia Impact Questionnaire; HADS, Hospital Anxiety and Depression Scale; POMS, Profile of Mood States;

✓ A study result is available for inclusion in the synthesis

X No study result is available for inclusion, (probably) because the P value, magnitude or direction of the results generated were considered unfavourable by the study investigators

-- No study result is available for inclusion, (probably) because the outcome was not assessed, or for a reason unrelated to the P value, magnitude, or direction of the results

? No study result is available for inclusion, and it is unclear if the outcome was assessed in the study

[^] Study data was not able to be extracted (results prior to crossover not reported)

D10.2.3.1 Primary Comparison (vs placebo)

One RCT (Bell 2004) and one quasi RCT (Fisher 1988) comparing homeopathy with placebo in people with fibromyalgia were eligible for this comparison. Bell 2004 contributed data to 4 outcomes considered critical or important for this review. Fisher 1988 reported data for 2 critical outcome domains, however the data was not presented in an extractable form.

There was one study awaiting classification (25 participants) that compared homeopathy with placebo for fibromyalgia that could have contributed data to 2 of the outcome domains critical for this review (fatigue and health-related quality of life). Complete results from this study were not available for inclusion, (probably) because the p-value, magnitude or direction of the results generated were considered unfavourable by the study investigators.

There were no ongoing studies that compared homeopathy with placebo.

Pain intensity

One study (53 participants) reported pain intensity measured using the McGill Pain Questionnaire (MPQ) at 3 months (one month prior to the end of treatment) (Bell 2004).

The MPQ is a 3-part assessment tool that measures sensory, affective, and evaluative pain dimensions (194). The questionnaire consists of 78 words, of which people choose those that best describe their experience of pain. Scores are tabulated by summing values associated with each word. Total scores range from zero, indicating no pain, to 78, indicating severe pain. Bell 2004 measured sensory pain, measured on a scale from 0 to 33 and affective pain measured on a scale from 0 to 12. The evaluative pain dimension was not reported. An MCID for the MPQ could not be identified.

Results from Bell 2004 suggested there is no difference in pain in the homeopathy group compared to the placebo group for either affective pain (MD -0.20; 95% CI -1.71, 1.31; $p = 0.80$) or sensory pain (MD 0.50; 95% CI -3.36, 4.36; $p = 0.80$) (*GRADE: Very Low*).

One study (Fisher 1988) measured pain intensity using a 10-cm Visual Analog Scale (VAS). The VAS is a validated tool used for the subjective measurement of pain on a continuous scale (usually from 0 to 100), with higher scores indicating worse levels of pain. The study authors reported dichotomised values, as the number of participants who improved, with a significant difference favouring the homeopathy group for combined pain and sleep scores ($p = 0.0052$). It is not clear if these data are before or after the crossover period.

Fatigue

One study (53 participants) reported fatigue using the Profile of Mood States (POMS) fatigue domain at 3 months (one month prior to the end of treatment) (Bell 2004).

POMS is a validated questionnaire designed to evaluate emotional wellbeing through 7 different mood domains, including fatigue (195). The POMS fatigue domain is scored on a scale from 0 to 28, where a higher score indicates worse fatigue. The MCID for this outcome could not be identified.

Results from this study suggested there was no difference in fatigue between the homeopathy treatment group and the placebo group (MD -3.40; 95% CI -7.47, 0.67; $p = 0.10$) (*GRADE: Very Low*).

Only 3 out of the 7 POMS domains were reported, raising concerns of selective reporting.

Health-related quality of life

One study (53 participants) reported health-related quality of life measured using the Global Health Rating at 3 months (one month prior to the end of treatment) (Bell 2004).

The Global Health Rating is scored from 3 questions relating to current health, health related to peers and health compared to 6-months ago. Each question is rated out of 5, resulting in an overall score ranging from 3 to 15, where a higher score indicates better health-related quality of life.

Results from Bell 2004 suggested there is no difference in health-related quality of life between the homeopathy group compared to the placebo group (MD 0.50; 95% CI -1.09, 2.09; $p = 0.54$) (*GRADE: Very Low*).

Emotional wellbeing

One study (53 participants) reported emotional wellbeing measured with the Profile of Mood States (POMS) at 3 months (one month prior to the end of treatment) (Bell 2004).

POMS is a validated questionnaire designed to evaluate emotional wellbeing through 7 different mood domains, including fatigue and depression (195). Only 3 out of the 7 POMS domains for emotional wellbeing were reported, raising concerns of selective reporting.

Results from Bell 2004 suggested there is no difference in emotional wellbeing between the homeopathy group and the placebo group for depression (MD -0.80; 95% CI -6.16, 4.56; $p = 0.77$) (*GRADE: Very Low*).

D10.2.3.2 Secondary Comparison (vs 'inactive' control)

One RCT (Relton 2009) (36 participants) comparing homeopathy with inactive control (no intervention) in people with fibromyalgia was eligible for this comparison and contributed data to 4 outcome domains considered critical for this review. All participants were allowed to continue with their usual care, which consisted of one or more of the following: physiotherapy, aerobic exercise, analgesics, non-steroidal anti-inflammatory drugs, or anti-depressants.

There were no studies awaiting classification and no ongoing studies that compared homeopathy to inactive control.

Pain intensity

One study (36 participants) reported pain intensity measured using the McGill Pain Questionnaire at end of treatment (22 months) (Relton 2009).

The McGill Pain Questionnaire is a 3-part assessment tool that measures sensory, affective, and evaluative pain dimensions (194). The questionnaire consists of 78 words, of which people choose those that best describe their experience of pain. Scores are tabulated by summing values associated with each word. Total scores range from zero, indicating no pain, to 78, indicating severe pain. Sensory pain is measured on a scale of 0 to 33, affective pain on a scale of 0 to 12, and evaluative pain is measured using a visual analogue scale (VAS) ranging from 0-100, where a higher score indicated worse pain.

Results from Relton 2009 suggested there is no difference in pain intensity between the homeopathy control groups for affective pain (MD -2.00; 95% CI -4.34, 0.34; $p = 0.09$); sensory pain (MD -2.90; 95% CI -8.94, 3.14; $p = 0.35$); or pain intensity (MD -14.00; 95% CI -28.37, 0.37; $p = 0.06$) (*GRADE: Low*).

Fatigue

One study (36 participants) reported fatigue using the Fibromyalgia Impact Questionnaire (FIQ) fatigue domain at the end of treatment (22 weeks) (Relton 2009). The FIQ is an evaluation tool developed to measure several key domains relating to function, impact and symptoms for fibromyalgia patients (196). The FIQ fatigue domain is measured on a scale from 0-10, where a higher score indicates worse fatigue. Results from Relton 2009 showed no difference in fatigue between the homeopathy group and the control group (MD -1.10, 95% CI: -2.44, 0.24) (*GRADE: Low*).

Health-related quality of life

One study (36 participants) reported health-related quality of life using the Fibromyalgia Impact Questionnaire (FIQ) at the end of treatment (22 weeks) (Relton 2009). The FIQ is an evaluation tool developed to measure quality of life through several key domains for fibromyalgia patients, including function, overall impact and symptoms (196). Total FIQ scores are measured on a scale from 0-100, where a higher score indicates worse overall quality of life. A 14-point change in the FIQ total score is considered clinically relevant (197). Results from this study showed there was no statistically significant difference in health-related quality of life between the homeopathy group and the control group (MD -10.30, 95% CI: -23.93, 3.33). (*GRADE: Low*)

Emotional wellbeing

One study (36 participants) reported emotional wellbeing measured using the Hospital Anxiety and Depression Scale (HADS) at the end of treatment (22 weeks) (Relton 2009). The HADS is a 14-question self-assessment tool that aims to measure depression and anxiety levels in medical patients (198). Emotional wellbeing is measured using total anxiety and depression scores on a scale of 0 to 42, where a higher score indicates worse emotional wellbeing. Results from Relton 2009 showed there was no statistically significant difference in emotional wellbeing between the homeopathy group and control group (MD -3.10, 95% CI: -8.85, 2.65). (*GRADE: Low*)

D10.2.3.3 Tertiary Comparison (vs other)

There were no RCTS comparing homeopathy with 'other' interventions in people with fibromyalgia that were eligible for this comparison.

Appendix E Risk of bias forms

This appendix (see attachment E) documents the risk of bias judgements made on studies that met the prespecified inclusion criteria for a systematic review on the effect of homeopathy for preventing and treating any health condition.

The risk of bias of included RCTs was assessed using the Revised Cochrane Risk of Bias tool v2.0 (199, 200).

Appendix E lists the included RCTs and quasi RCTs (for priority populations) in order of ICD-11 category. Studies within the ICD-11 category are then ordered by the prioritised condition and listed alphabetically. For each study there are 2 columns: column one is the judgement applied to each signalling question associated with each risk of bias domain (answered as yes, partial yes, no, partial no, no information or not applicable); column 2 is a comment that briefly explains the reasoning that underpins the judgement.

(see separate spreadsheet – Appendix E1-RoB)

Appendix F Characteristics of included studies

This appendix documents the data extracted from studies that met the prespecified inclusion criteria for a systematic review on the effect of homeopathy for preventing and treating any health condition and were conducted in populations prioritised for inclusion in the evidence synthesis.

All extracted data is presented, including that which was not synthesised in the main report.

F1 Study details

(see separate spreadsheet)

Appendix F1 (see attachment F1) lists the characteristics of each included study (for priority populations) in order of ICD-11 category. Studies within the ICD-11 category are then ordered by the prioritised condition and listed alphabetically.

For each study, the data extraction has included (but was not limited to) the following characteristics: study design, year conducted, setting and location, participant inclusion criteria, intervention and comparator characteristics (including number of treatment sessions, program duration, co-interventions), outcomes (including measurement method and timing), and funding sources.

Outcome domains and measures considered critical or important for inclusion in the review are highlighted with a blue box. Conversely, outcome domains and measures that were of limited importance are not highlighted.

F2 Supplementary outcome data

(see separate spreadsheet)

Appendix F2 (see separate attachment) lists the data extracted for critical or important outcomes identified in each included study (for priority populations) in order of ICD-11 category. Studies within the ICD-11 category are then ordered by the prioritised condition. Within each sheet, studies are listed by comparison (homeopathy vs 'inactive' control or homeopathy vs 'other') with the study results per outcome reported (critical or important outcome measures) that includes (but is not limited to) the following: outcome domain, timing, outcome measure, measure details, number of included participants, point estimates, p-value, direction of effect.

Data extracted is that reported by the study authors at the end of treatment (where possible) with footnotes included if further explanation was required (e.g. authors do not provide end-of treatment results therefore the mean change from baseline data are reported). The final column lists the risk of bias assessment for that outcome as made by the review authors (see [Appendix E](#)).

Appendix G Differences between protocol & review

G1 Methods not implemented

There were some methods that were not implemented in the review relating to the following sections:

Assessment of bias within NRSIs

Given NRSIs were not included in the evidence evaluation, the ROBINS-I risk of bias tool was not used.

Studies identified in the literature search

It was intended that, if a study did not contain the required PICO information for a decision to be made regarding its eligibility, the information would be sought from the study's authors through an open-ended request. Given time and resource constraints, we did not contact authors for additional information regarding eligibility criteria.

Requests for data

Eligible primary studies not published in English, ongoing trials and studies published as conference abstracts with incomplete results were identified for inclusion and listed as either 'Ongoing' or within the 'Studies Awaiting Classification'. It was intended that study authors would be contacted through an open-ended request for further information, and, if available, the study would be included in the evidence appraisal. Given time and resource constraints, we did not contact study authors for additional information regarding missing data.

Risk of reporting bias across studies

To assess potential bias due to 'non-reporting', it was intended that funnel plots (of effect estimates against their standard errors) would be generated in RevMan 5.4; with visual inspection of the funnel plot being used to look for evidence of asymmetry (suggesting small-study effects or missing results). Other possible reasons for funnel plot asymmetry were to be considered at this time (e.g. poor methodological quality, true heterogeneity, chance) (201). As there were less than ten RCTs included for any given PICO, inspection of funnel plots for patterns of asymmetry was not performed.

Quantitative synthesis

The NTWC could request that data comparing Homeopathy with 'other' (active) intervention be synthesised (prior to provision of the first draft evaluation report), where:

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

No such cases were identified or requested.

Subgroup analyses and investigations of heterogeneity

We did not plan to undertake any subgroup analyses of subsets of participants within or across studies. The planned subgroup analysis was to explore possible sources of heterogeneity relating to the intervention (i.e., individualised and non-individualised); however investigations relating to heterogeneity was not implemented as there were too few studies per analyses.

In particular, to allow for potential subgroup analyses, studies were to be stratified based on whether the participants received individualised or non-individualised homeopathy; however, there were few conditions for which there were two or more studies available for a comparison that also included studies with a different mode of intervention (individualised or non-individualised). We therefore did not stratify according to the intervention.

G2 Changes from protocol

There were some differences between the protocol and review relating to the following sections:

Nonrandomised interventional studies

The populations for eligible NRSIs were to be specified after initial screening of RCTs for prioritised populations (see Section 3.1.2 and Flowchart 2). The search strategy was to be implemented in 2 phases:

- All eligible RCTs will be found using the search strategy outlined in Appendix B. Independently, populations will be prioritised by the NTWC and NTREAP as outlined in Section 3.1.2 and Flowchart 2.
- If there are no RCTs identified for a priority population, an additional search for NRSIs in that population will be conducted. This second search will use the search strategy listed in Appendix B, augmented with population-specific search strings. The additional population filters will be developed and tested using the process described above and approved by the NTWC prior to implementation.

There was one condition initially listed as priority for which no RCT evidence was found (inflammatory bowel disease). We did not use the search strategy specified for nonrandomised studies, as the population specific search string (PubMed strings outlined below) found no evidence, without the need to applying study design filters.

("homeopathy s"[All Fields] OR "homoeopathy"[All Fields] OR "homeopathy"[MeSH Terms] OR "homeopathy"[All Fields]) AND ("colitis, ulcerative"[MeSH Terms] OR ("colitis"[All Fields] AND "ulcerative"[All Fields]) OR "ulcerative colitis"[All Fields] OR ("ulcerative"[All Fields] AND "colitis"[All Fields]) OR ("inflammatories"[All Fields] OR "inflammatory"[All Fields]) AND ("bowel s"[All Fields] OR "bowell"[All Fields] OR "intestines"[MeSH Terms] OR "intestines"[All Fields] OR "bowel"[All Fields] OR "bowels"[All Fields])) OR ("crohn disease"[MeSH Terms] OR ("crohn"[All Fields] AND "disease"[All Fields]) OR "crohn disease"[All Fields] OR "crohn s disease"[All Fields]))

Outcome measures and timepoints of interest

It was intended that outcomes reported at different timepoints were to be grouped and considered as either: short term, intermediate term, long-term (or not specified); with the NTWC to decide during outcome prioritisation as to whether evidence reported at multiple timepoints would be considered critical or important for decision-making (to be considered and reported separately). During the preliminary data extraction (and prior to outcome prioritisation), it became apparent that very few studies reported anything beyond baseline and end of treatment scores (i.e. there was minimal reporting of mid-treatment or follow-up results beyond completion of the treatment phase).

A pragmatic decision was therefore made to maximise the available data eligible for inclusion, with 'end-of-treatment' outcomes being the sole timepoint of interest to be considered in the evidence synthesis (unless there was good rationale for selecting an alternative timeframe).

Appendix H How comments from methodological review were addressed

Methodological review (or peer review) was conducted to appraise the methodological quality and assess the appropriateness of reporting for this systematic review (including appendices).

For reporting, the methodological review assessed the systematic review against the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Checklist (2020) and where applicable, the MECIR (Methodological Expectations of Cochrane Intervention Reviews) manual.

The ROBIS (Risk of Bias in Systematic Reviews) tool was used to assess the methodological quality of the systematic review, to ensure it was designed and conducted in accordance with:

- NHMRC's Developing your Guideline module in NHMRC's Guidelines for Guidelines Handbook
- Cochrane Handbook for Systematic Reviews of Interventions (updated 2022)
- GRADE guidance and GRADE working group criteria for determining whether the GRADE approach was used (GRADE handbook).

The ROBIS assessment included specification and application of criteria for considering studies for the review and synthesis, search methods, data extraction and analysis, assessment of risk of bias of studies, assessment of the certainty of evidence using GRADE, and the interpretation and summary of findings.

The systematic review (including appendices) has been updated to reflect the amendments suggested by methodological review and NHMRC's Natural Therapies Working Committee, where appropriate. In summary, updates included additional information and/ or clarification of the Plain Language Summary, Executive Summary, Results sections and Appendices, including:

- Edits to summary of findings tables including clarification of data reported, providing additional information within the footnotes and clarification of thresholds used to assess imprecision for some conditions.
- Clarification about the prioritisation process was provided in Appendix A6.

A detailed record of responses to all comments indicating changes that were made was provided to NHMRC together with the amended Report and Appendices documents.

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. Methodological Expectations of Cochrane Intervention Reviews. London: Cochrane; 2019 [cited 2019]. Available from: <https://community.cochrane.org/mecir-manual>.
3. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. 10.1371/journal.pmed.1000097
4. Churchill R, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, et al. Standards for the reporting of new Cochrane Intervention Reviews. 2019 [cited 4 June 2020]. In: Methodological Expectations of Cochrane Intervention Reviews [Internet]. London: Cochrane, [cited 4 June 2020]. Available from: <https://community.cochrane.org/mecir-manual/standards-reporting-new-cochrane-intervention-reviews-r1-r109>.
5. 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. [Accessed; (Updated October 2013). Available from: <https://gdt.gradeapro.org/app/handbook/handbook.html>].
6. Aabel S, Laerum E, Dolvik S, Djupesland P. Is homeopathic 'immunotherapy' effective? A double-blind, placebo-controlled trial with the isopathic remedy Betula 30c for patients with birch pollen allergy. *British Homeopathic Journal*. 2000;89(4):161-8. <http://dx.doi.org/10.1054/homp.1999.0430>
7. Aabel S. No beneficial effect of isopathic prophylactic treatment for birch pollen allergy during a low-pollen season: A double-blind, placebo-controlled clinical trial of homeopathic Betula 30c. *British Homeopathic Journal*. 2000;89(4):169-73. <http://dx.doi.org/10.1054/homp.1999.0440>
8. Aabel S. Prophylactic and acute treatment with the homeopathic medicine Betula 30c for birch pollen allergy: A double-blind, randomized, placebo-controlled study of consistency of VAS responses. *British Homeopathic Journal*. 2001;90(2):73-8. <http://dx.doi.org/10.1054/homp.1999.0471>
9. Kim LS, Riedlinger JE, Baldwin CM, Hilli L, Khalsa SV, Messer SA, et al. Treatment of seasonal allergic rhinitis using homeopathic preparation of common allergens in the southwest region of the US: a randomized, controlled clinical trial. *Ann Pharmacother*. 2005;39(4):617-24. 10.1345/aph.1E387
10. Liu L-L, Wan K-S, Cheng C-F, Tsai M-H, Wu Y-L, Wu W-F. Effectiveness of MORA electronic homeopathic copies of remedies for allergic rhinitis: A short-term, randomized, placebo-controlled PILOT study. *European Journal of Integrative Medicine*. 2013;5(2):119-25. 10.1016/j.eujim.2012.10.003
11. Naidoo P, Pellow J. A randomized placebo-controlled pilot study of Cat saliva 9cH and Histaminum 9cH in cat allergic adults. *Homeopathy*. 2013;102(2):123-9. 10.1016/j.homp.2013.02.007
12. Reilly D, Taylor M. A double blind placebo controlled study model for assessing homoeopathy using homoeopathic mixed grass pollens 30C in hay fever. *Comm Br Homoeopath Res Grp*. 1984;11:67-70.
13. Reilly DT, Taylor MA, McSharry C, Aitchison T. Is homoeopathy a placebo response? Controlled trial of homoeopathic potency, with pollen in hayfever as model. *Lancet*. 1986;2(8512):881-6.
14. Taylor MA, Reilly D, Llewellyn-Jones RH, McSharry C, Aitchison TC. Randomised controlled trial of homeopathy versus placebo in perennial allergic rhinitis with overview of four trial series...including commentary by Lancaster T and Vickers A [corrected] [published erratum appears in BMJ 2001 Feb 3; 322(7281): 282]. *BMJ: British Medical Journal (International Edition)*. 2000;321(7259):471-6. 10.1136/bmj.321.7259.471
15. Taylor MA. Erratum: Randomised controlled trial of homoeopathy versus placebo in perennial allergic rhinitis with overview of four trial series (British Medical Journal (Aug. 19-26) (471-476)). *British Medical Journal*. 2000;321(7263):733.
16. Taylor MA. Randomized, controlled trial of homeopathy versus placebo in perennial allergic rhinitis with overview of four trial series. *Indian journal of pediatrics*. 2002;69(8):674.
17. Wiesenauer M, Lüdtke R. The treatment of pollinosis with Galphimia glauca D4 - a randomized placebo-controlled double-blind clinical trial. *Phytomedicine*. 1995;2(1):3-6. 10.1016/s0944-7113(11)80041-3

18. Ellis AK, Soliman M, Steacy L, Boulay M, Boulet LP, Keith PK, et al. The Allergic Rhinitis - Clinical Investigator Collaborative (AR-CIC): nasal allergen challenge protocol optimization for studying AR pathophysiology and evaluating novel therapies. *Allergy Asthma Clin Immunol*. 2015;11(1):16. 10.1186/s13223-015-0082-0
19. American Thoracic Society. Rhinoconjunctivitis Quality of Life Questionnaire (RQLQ) n.d. [Available from: <https://www.thoracic.org/members/assemblies/assemblies/srn/questionnaires/rqlq.php>.
20. Carello R, Ricottini L, Miranda V, Panei P, Rocchi L, Arcieri R, et al. Long-term treatment with low-dose medicine in chronic childhood eczema: a double-blind two-stage randomized control trial. *Ital J Pediatr*. 2017;43(1):78. 10.1186/s13052-017-0393-5
21. Dey S, Shaikh AR, Agrawal E, Gautam AK, Karuppusamy A, Sadhukhan S, et al. Efficacy of Individualized Homeopathic Medicines in the Treatment of Atopic Dermatitis in Adults: A Double-Blind, Randomized, Placebo-Controlled, Preliminary Trial. *Complementary Medicine Research*. 2022;29(1):17-26. <https://dx.doi.org/10.1159/000516026>
22. Fisher P, McCarney R, Hasford C, Vickers A. Evaluation of specific and non-specific effects in homeopathy: Feasibility study for a randomised trial. *Homeopathy: the Journal of the Faculty of Homeopathy*. 2006;95(4):215-22. <https://dx.doi.org/10.1016/j.homp.2006.07.006>
23. Vickers A. Evaluation of specific and non-specific effects in homeopathy: Feasibility study for a randomised trial. *British Homeopathic Journal*. 2000;89(SUPPL. 1):S48-S9. <https://dx.doi.org/10.1054/homp.1999.0388>
24. Oakley A. SCORAD: DermNet; 2009 [Available from: <https://dermnetnz.org/topics/scorad#:~:text=SCORAD%20is%20a%20clinical%20tool,the%20treatment%20has%20been%20effective>.
25. Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology*. 2015;230(1):27-33. 10.1159/000365390
26. Australasian Society of Clinical Immunology and Allergy. Eczema (Atopic Dermatitis) Frequently Asked Questions (FAQ) March 2023 [November 2023]. Available from: <https://www.allergy.org.au/patients/skin-allergy/eczema?highlight=WyJlY3plbWEiXQ==>.
27. Jacobs J, Springer DA, Crothers D. Homeopathic treatment of acute otitis media in children: A preliminary randomized placebo-controlled trial. *Pediatric Infectious Disease Journal*. 2001;20(2):177-83. <https://dx.doi.org/10.1097/00006454-200102000-00012>
28. Jacobs J, Springer DA, Crothers D. Homeopathic treatment of otitis media in children: a preliminary randomized controlled trial with placebo. *Homeopatia mex*. 2002;71(618):109-21.
29. Pedrero-Escalas MF, Jimenez-Antolin J, Lassaletta L, Diaz-Saez G, Gavilan J. Hospital clinical trial: Homeopathy (Agraphis nutans 5CH, Thuya occidentalis 5CH, Kalium muriaticum 9CH and Arsenicum iodatum 9CH) as adjuvant, in children with otitis media with effusion. *Int J Pediatr Otorhinolaryngol*. 2016;88:217-23. 10.1016/j.ijporl.2016.06.039
30. Harrison H, Fixsen A, Vickers A. A randomized comparison of homoeopathic and standard care for the treatment of glue ear in children. *Complementary Therapies in Medicine*. 1999;7(3):132-5. <https://dx.doi.org/10.1016/S0965-2299%2899%2980120-4>
31. Taylor JA, Jacobs J. Homeopathic ear drops as an adjunct to standard therapy in children with acute otitis media. *Homeopathy*. 2011;100(3):109-15. 10.1016/j.homp.2011.03.002
32. Taylor JA, Jacobs J. Homeopathic Ear Drops as an Adjunct in Reducing Antibiotic Usage in Children With Acute Otitis Media. *Glob Pediatr Health*. 2014;1:2333794X14559395. 10.1177/2333794X14559395
33. Taylor JA, Jacobs J. Homeopathic treatment of respiratory illnesses in children: Results from two randomized trials. *Homeopathy: the Journal of the Faculty of Homeopathy*. 2016;105(1):15. <https://dx.doi.org/10.1016/j.homp.2015.12.006>
34. Sinha MN, Siddiqui VA, Nayak C, Singh V, Dixit R, Dewan D, et al. Randomized controlled pilot study to compare Homeopathy and Conventional therapy in Acute Otitis Media. *Homeopathy*. 2012;101(1):5-12. 10.1016/j.homp.2011.08.003
35. de Lange de Klerk ESM. Effects of homeopathic medicines on children with recurrent upper respiratory tract infections. *HomInt R&D Newsletter*. 1993.

36. De Lange de Klerk ESM, Blommers J, Kruik DJ, Bezemer PD, Feenstra L. Effect of homoeopathic: Medicines on daily burden of symptoms in children with recurrent upper respiratory tract infections. *British Medical Journal*. 1994;309(6965):1329-32.
37. De Lange De Klerk ES, Blommers J, Kuik DJ, Feenstra L, Bezemer PD. Effects of individually chosen homeopathic medicines on recurrent URTI in children. A clinical trial - I study methodology. *British homoeopathic journal*. 1996(85):4-14.
38. Furuta SE, Weckx LLM, Figueiredo CR. Randomized, double-blind trial on the efficacy of homeopathic treatment in children with recurrent tonsillitis. *Rev homeopatia (Sao Paulo)*. 2017;80(3/4):136-41.
39. Steinsbekk A, Bentzen N, Fonnebo V, Lewith GT. Randomized controlled trials on treatment by homeopaths and self-treatment with homeopathic medicines: design and protocol. *J Altern Complement Med*. 2004;10(6):1027-32. 10.1089/acm.2004.10.1027
40. Steinsbekk A, Bentzen N, Fonnebo V, Lewith G. Self treatment with one of three self selected, ultramolecular homeopathic medicines for the prevention of upper respiratory tract infections in children. A double-blind randomized placebo controlled trial. *Br J Clin Pharmacol*. 2005;59(4):447-55. 10.1111/j.1365-2125.2005.02336.x
41. Steinsbekk A, Fonnebo V, Lewith G, Bentzen N. Homeopathic care for the prevention of upper respiratory tract infections in children: a pragmatic, randomised, controlled trial comparing individualised homeopathic care and waiting-list controls. *Complement Ther Med*. 2005;13(4):231-8. 10.1016/j.ctim.2005.06.007
42. Steinsbekk A, Lewith G, Fonnebo V, Bentzen N. An exploratory study of the contextual effect of homeopathic care. A randomised controlled trial of homeopathic care vs. self-prescribed homeopathic medicine in the prevention of upper respiratory tract infections in children. *Prev Med*. 2007;45(4):274-9; discussion 80-1. 10.1016/j.ypmed.2007.02.004
43. Palm J, Kishchuk VV, Ulied A, Fernandez JP, De Jaegere S, Jong MC, et al. Effectiveness of an add-on treatment with the homeopathic medication SilAto-5-90 in recurrent tonsillitis: An international, pragmatic, randomized, controlled clinical trial. *Complementary Therapies in Clinical Practice*. 2017;28:181-91. <https://dx.doi.org/10.1016/j.ctcp.2017.05.005>
44. Pannek J, Pannek-Rademacher S, Jus MS, Wollner J, Krebs J. Usefulness of classical homeopathy for the prophylaxis of recurrent urinary tract infections in individuals with chronic neurogenic lower urinary tract dysfunction. *J Spinal Cord Med*. 2019;42(4):453-9. 10.1080/10790268.2018.1440692
45. Witt A, Kaufmann U, Bitschnau M, Tempfer C, Ozbal A, Haytouglu E, et al. Monthly itraconazole versus classic homeopathy for the treatment of recurrent vulvovaginal candidiasis: a randomised trial. *BJOG*. 2009;116(11):1499-505. 10.1111/j.1471-0528.2009.02262.x
46. Baker DG, Myers SP, Howden I, Brooks L. The effects of homeopathic Argentum nitricum on test anxiety. *Complementary Therapies in Medicine*. 2003;11(2):65-71. <https://dx.doi.org/10.1016/S0965-2299%2803%2900059-1>
47. Bonne O, Shemer Y, Goral Y, Katz M, Shalev AY. A randomized, double-blind, placebo-controlled study of classical homeopathy in generalized anxiety disorder. *Journal of Clinical Psychiatry*. 2003;64(3):282-7. <http://dx.doi.org/10.4088/JCP.v64n0309>
48. Dimpfel. Psychophysiological Effectiveness of Calmvalera Hevert Tablets as Measured by EnkephaloVision in Anxious Subjects during Audio-Visual Cognitive and Emotional Challenges: A Double-Blind, Randomized, Placebo-Controlled, 2-Armed, Phase IV Study in Parallel Design. 2016.
49. Nct. Reducing Children's Anxiety Using Homeopathic Remedy Before Dental Treatment. <https://clinicaltrials.gov/show/NCT01658059>. 2012.
50. Parewa M, Burman AS, Brahma A, Rutten L, Sadhukhan S, Misra P, et al. Individualized Homeopathic Medicines in the Treatment of Generalized Anxiety Disorder: A Double-Blind, Randomized, Placebo-Controlled, Pilot Trial. *Complement Med Res*. 2021;28(5):407-18. 10.1159/000514524
51. Rickels K, Zaninelli R, McCafferty J, Bellew K, Iyengar M, Sheehan D. Paroxetine treatment of generalized anxiety disorder: a double-blind, placebo-controlled study. *Am J Psychiatry*. 2003;160(4):749-56. 10.1176/appi.ajp.160.4.749
52. Benson J, El-Zahhar N. Further refinement and validation of the revised test anxiety scale. *Structural Equation Modeling: A Multidisciplinary Journal*. 1994;1(3):203-21. 10.1080/10705519409539975

53. Michael PH, Martin P. Estimates of the minimal important difference to evaluate the clinical significance of antidepressants in the acute treatment of moderate-to-severe depression. *BMJ Evidence-Based Medicine*. 2022;27(2):69. 10.1136/bmjebm-2020-111600
54. Aripiprazole (Abilify): Depression, Major Depressive Disorder (MDD). Ottawa2016 [cited 2016]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK409740/table/T25/>.
55. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med*. 1983;13(3):595-605.
56. Adler UC, Kruger S, Teut M, Ludtke R, Bartsch I, Schutzler L, et al. Homeopathy for depression--DEP-HOM: study protocol for a randomized, partially double-blind, placebo controlled, four armed study. *Trials*. 2011;12:43. 10.1186/1745-6215-12-43
57. Adler UC, Kruger S, Teut M, Ludtke R, Schutzler L, Martins F, et al. Homeopathy for depression: a randomized, partially double-blind, placebo-controlled, four-armed study (DEP-HOM). *PLoS One*. 2013;8(9):e74537. 10.1371/journal.pone.0074537
58. Katz T, Fisher P, Katz A, Davidson J, Feder G. The feasibility of a randomised, placebo-controlled clinical trial of homeopathic treatment of depression in general practice. *Homeopathy*. 2005;94(3):145-52. 10.1016/j.homp.2005.04.002
59. Viksveen P, Relton C. Depression treated by homeopaths: a study protocol for a pragmatic cohort multiple randomised controlled trial. *Homeopathy*. 2014;103(2):147-52. 10.1016/j.homp.2014.01.004
60. Viksveen P. Homeopathy in self-reported depression: A pragmatic randomised controlled trial. *Homeopathy: the Journal of the Faculty of Homeopathy*. 2016;105(1):24. <https://dx.doi.org/10.1016/j.homp.2015.12.006>
61. Viksveen P, Relton C, Nicholl J. Depressed patients treated by homeopaths: a randomised controlled trial using the "cohort multiple randomised controlled trial" (cmRCT) design. *Trials*. 2017;18(1):299. 10.1186/s13063-017-2040-2
62. Adler UC, Paiva NM, Cesar AT, Adler MS, Molina A, Padula AE, et al. Homeopathic Individualized Q-potencies versus Fluoxetine for Moderate to Severe Depression: Double-blind, Randomized Non-inferiority Trial. *Evidence-based complementary and alternative medicine : eCAM*. 2009(nep114):[8]-[].
63. Adler UC, Paiva NM, Cesar AT, Adler MS, Molina A, Padula AE, et al. Homeopathic Individualized Q-Potencies versus Fluoxetine for Moderate to Severe Depression: Double-Blind, Randomized Non-Inferiority Trial. *Evid Based Complement Alternat Med*. 2011;2011:520182. 10.1093/ecam/nep114
64. Calil HM, Adler UC, Paiva NMP, Cesar AT, Adler MS, Molina A, et al. Homeopathic individualized Q-potencies versus fluoxetine for moderate to severe depression: Double-blind, randomized non-inferiority trial. *Evidence-based Complementary and Alternative Medicine*. 2011;2011 (no pagination). <https://dx.doi.org/10.1093/ecam/nep114>
65. User's manual for the SF-36v2 survey. 3rd ed. Lincoln RD, editor: Quality Metric Incorporated; 2011.
66. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13. 10.1046/j.1525-1497.2001.016009606.x
67. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care*. 2004;42(12):1194-201. 10.1097/00005650-200412000-00006
68. Dhawale K, Katawala M, Tambitkar N, Tamboli M, Tamboli P. Use of homoeopathic remedies in the management of learning disabilities. *Indian Journal of Research in Homoeopathy*. 2014;8:87. 10.4103/0974-7168.135641
69. Frei H, Everts R, von Ammon K, Kaufmann F, Walther D, Hsu-Schmitz SF, et al. Homeopathic treatment of children with attention deficit hyperactivity disorder: a randomised, double blind, placebo controlled crossover trial. *Eur J Pediatr*. 2005;164(12):758-67. 10.1007/s00431-005-1735-7
70. Frei H, Everts R, von AK, Thurneysen A. Homoopathische Behandlung von hyperaktiven Kindern: Ergebnisse einer randomisierten, placebo-kontrollierten Doppelblindstudie mit Crossover. *Zeitschrift für Klassische Homoopathie*. 2006;50(1):5-12.
71. Frei H, Everts R, von Ammon K, Kaufmann F, Walther D, Schmitz SFH, et al. Randomised controlled trials of homeopathy in hyperactive children: treatment procedure leads to an unconventional study design. Experience with open-label homeopathic treatment preceding the Swiss ADHD placebo

- controlled, randomised, double-blind, cross-over trial. *Homeopathy: the Journal of the Faculty of Homeopathy*. 2007;96(1):35-41. <https://dx.doi.org/10.1016/j.homp.2006.11.004>
72. von Ammon K, Frei H, Kaufmann F, Schmitz SH, Steinlin M, Thurneysen A. Homoeopathic RCT embedded in an observational study of children with attention deficit hyperactivity disorder -- a successful model of whole-systems CAM research...14th Annual Symposium on Complementary Health Care, 11th to 13th December 2007, University of Exeter, UK. *Focus on Alternative & Complementary Therapies*. 2007;12:5-.
73. Klaus VA, Ursula S, Heiner F, Sabine K, Andre T, Martin FE. Classical homeopathy helps hyperactive children-a 10-year follow-up of homeopathic and integrated medical treatment in children suffering from attention deficit disorder with and without hyperactivity. *European Journal of Integrative Medicine*. 2012;1:73-4. <https://dx.doi.org/10.1016/j.eujim.2012.07.645>
74. Jacobs J, Williams AL, Girard C, Njike VY, Katz D. Homeopathy for attention-deficit/hyperactivity disorder: a pilot randomized-controlled trial. *J Altern Complement Med*. 2005;11(5):799-806. 10.1089/acm.2005.11.799
75. Lamont J. Homoeopathic treatment of attention deficit hyperactivity disorder: a controlled study. *Br Homoeopath J*. 1997;86(10):196-200.
76. Ctri. A pilot study to see the effectiveness of homoeopathic therapy in attention Deficit Hyperactive Disorder of children. <https://trialsearchwho.int/Trial2.aspx?TrialID=CTRI/2011/12/002305>. 2011.
77. Strauss L. The efficacy of a homeopathic preparation in the management of attention deficit hyperactivity disorder. *Biomed Ther*. 2000;18(2):197-201.
78. Fibert P. Rationale for a pragmatic randomised controlled trial of the effectiveness of treatment by homeopaths for ADHD. *Homeopathy: the Journal of the Faculty of Homeopathy*. 2016;105(1):28-9. <https://dx.doi.org/10.1016/j.homp.2015.12.042>
79. Fibert P, Peasgood T, Relton C. Rethinking ADHD intervention trials: feasibility testing of two treatments and a methodology. *Eur J Pediatr*. 2019;178(7):983-93. 10.1007/s00431-019-03374-z
80. Fibert P, Relton C. Protocol for the Sheffield treatments for ADHD research (STAR) project: Testing the feasibility of the trial design. *ADHD Attention Deficit and Hyperactivity Disorders*. 2015;1:S39. <https://dx.doi.org/10.1007/s12402-015-0169-y>
81. Fibert P, Relton C. Preliminary feasibility and clinical results of a pilot study of treatment by homeopaths for children with adhd using the trials within cohorts (TwICs) design. *Homeopathy Conference: 3rd International Research Conference on "Cutting Edge Research in Homeopathy", HRI*. 2017;107(Supplement 1). <https://dx.doi.org/10.1055/s-0038-1632419>
82. Fibert P, Relton C. Preliminary Feasibility and Clinical Results of a Pilot Study of Treatment by Homeopaths for Children with ADHD using the Trials within Cohorts (TwICs) Design...HRI Malta 2017;Cutting Edge Research in Homeopathy: Presentation Abstracts. *Homeopathy: the Journal of the Faculty of Homeopathy*. 2018;107:55-78. 10.1055/s-0038-1632419
83. Fibert P, Relton C, Peasgood T, Daley D. Protocol for the STAR (Sheffield Treatments for ADHD) project: an internal pilot study assessing the feasibility of the Trials within Cohorts (TwICs) design to test the effectiveness of interventions for children with ADHD. *Pilot Feasibility Stud*. 2018;4:61. 10.1186/s40814-018-0250-3
84. Volpe RJ, Dupaul GJ. CHAPTER 12 - Assessment with Brief Behavior Rating Scales. In: Andrews JJW, Saklofske DH, Janzen HL, editors. *Handbook of Psychoeducational Assessment*. San Diego: Academic Press; 2001. p. 357-87.
85. Brulé D, Landau-Halpern B, Nastase V, Zemans M, Mitsakakis N, Boon H. A Randomized Three-Arm Double-Blind Placebo-Controlled Study of Homeopathic Treatment of Children and Youth with Attention-Deficit/Hyperactivity Disorder. *J Integr Complement Med*. 2023. 10.1089/jicm.2023.0043
86. The University of Sheffield. CHU9D - Measuring health and calculating QALYs for children and adolescents. [Available from: <https://licensing.sheffield.ac.uk/product/CHU-9D>.
87. Harrison CC, Solomon EM, Pellow J. The effect of a homeopathic complex on psychophysiological onset insomnia in males: a randomized pilot study. *Altern Ther Health Med*. 2013;19(5):38-43.
88. James M. Efficacy of individualized homeopathic treatment of insomnia: Double-blind, randomized, placebo-controlled clinical trial. *Complementary Therapies in Medicine*. 2019;43:53-9. <https://dx.doi.org/10.1016/j.ctim.2019.01.007>

89. Naude DF, Stephanie Couchman IM, Maharaj A. Chronic primary insomnia: efficacy of homeopathic simillimum. *Homeopathy*. 2010;99(1):63-8. 10.1016/j.homp.2009.11.001
90. Burkart J, Jong MC, Ilenkyo L, Kholodova I, Verwer C, Weber S, et al. Results of a randomised controlled trial with a homeopathic complex medicinal product in children with sleep disorders and restlessness. *Archives of Disease in Childhood*. 2017;102(Supplement 2):A100-A1. <https://dx.doi.org/10.1136/archdischild-2017-313273.260>
91. Jong C, Ilyenko L, Kholodova I, Verwer C, Burkart J, Weber S, et al. A comparative randomized controlled clinical trial on the effectiveness, safety, and tolerability of a homeopathic medicinal product in children with sleep disorders and restlessness. [Bulgarian]. *Pediatrica*. 2016;56(4):55-60. <https://dx.doi.org/10.1155/2016/9539030>
92. Jong MC, Ilyenko L, Kholodova I, Verwer C, Burkart J, Weber S, et al. A Comparative Randomized Controlled Clinical Trial on the Effectiveness, Safety, and Tolerability of a Homeopathic Medicinal Product in Children with Sleep Disorders and Restlessness. *Evid Based Complement Alternat Med*. 2016;2016(4):9539030. 10.1155/2016/9539030
93. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*. 2011;34(5):601-8. 10.1093/sleep/34.5.601
94. Morin CM, Azrin NH. Behavioral and cognitive treatments of geriatric insomnia. *J Consult Clin Psychol*. 1988;56(5):748-53. 10.1037//0022-006x.56.5.748
95. Rosch PJ. Insomnia, psychological assessment and management, Charles M. Morin, Guilford Press, New York, 1993. No. of pages: 238. Price: \$26.95. *Stress Medicine*. 1994;10(1):69-70. <https://doi.org/10.1002/smi.2460100113>
96. Cheuk DKL, Yeung WF, Chung KF, Wong V. Acupuncture for insomnia. *Cochrane Database of Systematic Reviews*. 2012(9). 10.1002/14651858.CD005472.pub3
97. Straumsheim P, Borchgrevink C, Mowinckel P, Kierulf H, Hafslund O. Homeopathic treatment of migraine: A double blind, placebo controlled trial of 68 patients. *British Homeopathic Journal*. 2000;89(1):4-7. <http://dx.doi.org/10.1054/homp.1999.0332>
98. Straumsheim PA, Borchgrevink C, Mowinckel P, Kierulf H, Hafslund O. Homeopathic treatment of migraine. A double-blind placebo controlled trial of 68 patients. *Dynamis (Granada, Spain)*. 1997;2.
99. Gaus W. Biometrische Aspekte der 'Munchener Kopfschmerzstudie'. *Allgemeine Homöopathische Zeitung*. 1997;242(6):245-9.
100. Gaus W, Walach H, Haag G. [The efficacy of classic homeopathic therapy in chronic headache. Study protocol.]. *Schmerz*. 1992;6(2):134-40. 10.1007/BF02528131
101. Walach H, Haeusler W, Lowes T, Mussbach D, Schamell U, Springer W, et al. Classical homeopathic treatment of chronic headaches. *Cephalalgia*. 1997;17(2):119-26. <https://dx.doi.org/10.1046/j.1468-2982.1997.1702119.x>
102. Walach H, Lowes T, Mussbach D, Schamell U, Springer W, Stritzl G, et al. The long-term effects of homeopathic treatment of chronic headaches: 1 Year follow up. *Cephalalgia*. 2000;20(9):835-7. <https://dx.doi.org/10.1046/j.1468-2982.2000.00129.x>
103. Whitmarsh TE, Coleston-Shields DM, Steiner TJ. Double-blind randomized placebo-controlled study of homeopathic prophylaxis of migraine. *Cephalalgia*. 1997;17(5):600-4. <https://dx.doi.org/10.1046/j.1468-2982.1997.1705600.x>
104. Lewith GT, Watkins AD, Hyland ME, Shaw S, Broomfield JA, Dolan G, et al. Use of ultramolecular potencies of allergen to treat asthmatic people allergic to house dust mite: Double blind randomised controlled clinical trial. *British Medical Journal*. 2002;324(7336):520-3.
105. Qutubuddin M, Murty Singh S, Nayak C, Koley M, Saha S. Efficacy of individualized homeopathy in bronchial asthma in adults: Double-blind, randomized, placebo-controlled, clinical trial in the context of usual care. *Advances in Integrative Medicine*. 2019;6(2):58-65. 10.1016/j.aimed.2018.08.005
106. Reilly D, Taylor M, Beattie N, Campbell J, McSharry C, Aitchison T, et al. Is evidence for homeopathy reproducible? *Lancet*. 1994;344(8937):1601-6.
107. Reilly D, Taylor M, Campbell J, Beattie N, McSharry C, Aitchison T, et al. Is homeopathy a placebo response? A controlled trial of Homeopathic Immunotherapy HIT in atopic asthma. *Complementary Therapies in Medicine*. 1993;1(Suppl 1):24-35.

108. White A, Slade P, Hunt C, Hart A, Ernst E. Individualised homeopathy as an adjunct in the treatment of childhood asthma: a randomised placebo controlled trial. *Thorax*. 2003;58(4):317-21. 10.1136/thorax.58.4.317
109. Thompson E, Shaw A, Montgomery A, Davies E, Sharpe D. A pilot randomised controlled trial of individualised homeopathy in children with severe asthma. Abstracts from the Developing Research Strategies conference, March 2007, Northampton, UK *Complementary Therapies in Medicine*. 2008;16(4):233-7. <https://doi.org/10.1016/j.ctim.2008.04.005>
110. Thompson EA, Shaw A, Nichol J, Hollinghurst S, Henderson AJ, Thompson T, et al. The feasibility of a pragmatic randomised controlled trial to compare usual care with usual care plus individualised homeopathy, in children requiring secondary care for asthma. *Homeopathy*. 2011;100(3):122-30. 10.1016/j.homp.2011.05.001
111. Topcu A, Lokke A, Eriksen L, Nielsen LP, Dahl R. Evaluating the effect on asthma quality of life of added reflexology or homeopathy to conventional asthma managementâan investigator-blinded, randomised, controlled parallel group study. *European Clinical Respiratory Journal*. 2020;7(1). 10.1080/20018525.2020.1793526
112. Topcu A, Lokke Ottesen A, Eriksen L, Nielsen LP, Dahl R. The impact of reflexology and homeopathy added to conventional asthma treatment on markers of airway inflammationÃ¢ randomised study. *European Clinical Respiratory Journal*. 2020;7(1). 10.1080/20018525.2020.1726153
113. Topcu A, Ottesen AL, Eriksen L, Nielsen LP, Dahl R. The Impact of Reflexology and Homeopathy Added to Conventional ASTHMA Treatment on Markers of Airway Inflammation - A Randomised Study. *European Respiratory Journal Conference: European Respiratory Society International Congress, ERS*. 2020;56(Supplement 64). <https://dx.doi.org/10.1183/13993003.congress-2020.2620>
114. Juniper EF, Svensson K, Mörk AC, Ståhl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. *Respir Med*. 2005;99(5):553-8. 10.1016/j.rmed.2004.10.008
115. Spirometry: Asthma Australia; [cited 2023 17 July]. Available from: <https://asthma.org.au/about-asthma/understanding-asthma/spirometry/>.
116. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. *Eur Respir J*. 2005;26(5):948-68. 10.1183/09031936.05.00035205
117. Bonini M, Paolo MD, Bagnasco D, Baiardini I, Braido F, Caminati M, et al. Minimal clinically important difference for asthma endpoints: an expert consensus report. *European Respiratory Review*. 2020;29(156):190137. 10.1183/16000617.0137-2019
118. Tepper RS, Wise RS, Covar R, Irvin CG, Kerckmar CM, Kraft M, et al. Asthma outcomes: pulmonary physiology. *J Allergy Clin Immunol*. 2012;129(3 Suppl):S65-87. 10.1016/j.jaci.2011.12.986
119. Bradley W DeVrieze, Pranav Modi, Giwa. AO. StatPearls (Internet). National Library of Medicine: StatPearls Publishing; 2002 [cited 2023 10 September]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK459325/>.
120. Hyland ME, Ley A, Fisher DW, Woodward V. Measurement of psychological distress in asthma and asthma management programmes. *Br J Clin Psychol*. 1995;34(4):601-11. 10.1111/j.2044-8260.1995.tb01494.x
121. Nishimura K, Hajiuro T, Oga T, Tsukino M, Sato S, Ikeda A. A comparison of two simple measures to evaluate the health status of asthmatics: the Asthma Bother Profile and the Airways Questionnaire 20. *J Asthma*. 2004;41(2):141-6. 10.1081/jas-120026070
122. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol*. 1994;47(1):81-7. 10.1016/0895-4356(94)90036-1
123. Jacobs J, Jimenez L, Gloyd S, Carares F, Gaitan M, Crothers D. Homoeopathic treatment of acute childhood diarrhoea. *Br Homoeopath J*. 1993;82(2):83-6.
124. Jacobs J, Jimenez LM, Malthouse S, Chapman E, Crothers D, Masuk M, et al. Homeopathic treatment of acute childhood diarrhea: results from a clinical trial in Nepal. *J Altern Complement Med*. 2000;6(2):131-9. 10.1089/acm.2000.6.131
125. Jacobs J, Fernandez EA, Merizalde B, Avila-Montes GA, Crothers D. The use of homeopathic combination remedy for dengue fever symptoms: a pilot RCT in Honduras. *Homeopathy*. 2007;96(1):22-6. 10.1016/j.homp.2006.10.004

126. Manoj P, Pawaskar N, Mundra P, Tamboli P, Kothare G. An approach to acute diarrhoeal disorders through sector and constitutional homeopathic treatment in tribal children attending balwadis. *Indian Journal of Research in Homoeopathy*. 2010;4(4):36-48. 10.53945/2320-7094.1763
127. Dossett ML, Mu L, Davis RB, Bell IR, Lembo AJ, Kaptchuk TJ, et al. Patient-Provider Interactions Affect Symptoms in Gastroesophageal Reflux Disease: A Pilot Randomized, Double-Blind, Placebo-Controlled Trial. *PLoS One*. 2015;10(9):e0136855. 10.1371/journal.pone.0136855
128. Paterson C, Ewings P, Brazier JE, Britten N. Treating dyspepsia with acupuncture and homeopathy: Reflections on a pilot study by researchers, practitioners and participants. *Complementary Therapies in Medicine*. 2003;11(2):78-84. <https://dx.doi.org/10.1016/S0965-2299%2803%2900056-6>
129. Raak C, Krueger P, Klement P, De Jaegere S, Weber S, Keller T, et al. Effectiveness of a Homeopathic Complex Remedy in Colic and Flatulence in Children: A randomized multicenter study. *Advances in Integrative Medicine*. 2019;6(Supplement 1):S57. <https://dx.doi.org/10.1016/j.aimed.2019.03.165>
130. Velanovich V. 25 Years of the GERD-HRQL symptom severity instrument: an assessment of published applications. *Surgical endoscopy*. 2023;37(1):255-65. 10.1007/s00464-022-09463-9
131. Ishaque S, Johnson JA, Vohra S. Individualized health-related quality of life instrument Measure Yourself Medical Outcome Profile (MYMOP) and its adaptations: a critical appraisal. *Quality of life research*. 2019;28(4):879-93. 10.1007/s11136-018-2046-6
132. Peckham EJ, Relton C, Raw J, Walters C, Thomas K, Smith C. A protocol for a trial of homeopathic treatment for irritable bowel syndrome. *BMC Complement Altern Med*. 2012;12:212. 10.1186/1472-6882-12-212
133. Peckham EJ, Relton C, Raw J, Walters C, Thomas K, Smith C, et al. Interim results of a randomised controlled trial of homeopathic treatment for irritable bowel syndrome. *Homeopathy*. 2014;103(3):172-7. <https://dx.doi.org/10.1016/j.homp.2014.05.001>
134. Das AD, Ghosh S, Palanisamy C, Guha N, Mandal S, Maiti S, et al. Efficacy of individualized homeopathic medicines in irritable bowel syndrome: A double-blind randomized, placebo-controlled trial. *EXPLORE*. 2023;19(4):519-27. <https://doi.org/10.1016/j.explore.2022.10.004>
135. Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. *Aliment Pharmacol Ther*. 1997;11(2):395-402. 10.1046/j.1365-2036.1997.142318000.x
136. Spiegel B, Camilleri M, Bolus R, Andresen V, Chey WD, Fehnel S, et al. Psychometric evaluation of patient-reported outcomes in irritable bowel syndrome randomized controlled trials: a Rome Foundation report. *Gastroenterology*. 2009;137(6):1944-53.e1-3. 10.1053/j.gastro.2009.08.047
137. Min Ho PY, Hu W, Lee YY, Gao C, Tan YZ, Cheen HH, et al. Health-related quality of life of patients with inflammatory bowel disease in Singapore. *Intest Res*. 2019;17(1):107-18. 10.5217/ir.2018.00099
138. Wiesenauer M. Research in homoeopathy: efficacy and tolerance of Mahonia aquifolium during treatment of psoriasis vulgaris. *Extracta Dermatologica*. 1992;16/12:23-31.
139. Bernstein S, Donsky H, Gulliver W, Hamilton D, Nobel S, Norman R. Treatment of mild to moderate psoriasis with Relieva, a Mahonia aquifolium extract--a double-blind, placebo-controlled study. *Am J Ther*. 2006;13(2):121-6. 10.1097/00045391-200603000-00007
140. Matza LS, Brazier JE, Stewart KD, Pinto L, Bender RH, Kircik L, et al. Developing a preference-based utility scoring algorithm for the Psoriasis Area Severity Index (PASI). *J Med Econ*. 2019;22(9):936-44. 10.1080/13696998.2019.1627362
141. Carlin CS, Feldman SR, Krueger JG, Menter A, Krueger GG. A 50% reduction in the Psoriasis Area and Severity Index (PASI 50) is a clinically significant endpoint in the assessment of psoriasis. *J Am Acad Dermatol*. 2004;50(6):859-66. 10.1016/j.jaad.2003.09.014
142. Koley M, Saha S, Ghosh S. A double-blind randomized placebo-controlled feasibility study evaluating individualized homeopathy in managing pain of knee osteoarthritis. *J Evid Based Complementary Altern Med*. 2015;20(3):186-91. 10.1177/2156587214568668
143. Shipley M, Berry H, Broster G, Jenkins M, Clover A, Williams I. Controlled trial of homoeopathic treatment of osteoarthritis. *Lancet*. 1983;1(8316):97-8.
144. Brien S, Lachance L, Lewith GT. Are the therapeutic effects of homeopathy attributed to the consultation, the homeopathic remedy, or both? A protocol for a future exploratory feasibility trial in

- patients with rheumatoid arthritis. *J Altern Complement Med*. 2004;10(3):499-502. 10.1089/1075553041323768
145. Brien S, Lachance L, Prescott P, McDermott C, Lewith G. Homeopathy has clinical benefits in rheumatoid arthritis patients that are attributable to the consultation process but not the homeopathic remedy: a randomized controlled clinical trial. *Rheumatology (Oxford)*. 2011;50(6):1070-82. 10.1093/rheumatology/keq234
146. Brien SB, Leydon GM, Lewith G. Homeopathy enables rheumatoid arthritis patients to cope with their chronic ill health: a qualitative study of patient's perceptions of the homeopathic consultation. *Patient Educ Couns*. 2012;89(3):507-16. 10.1016/j.pec.2011.11.008
147. Fisher P, Scott DL. A randomized controlled trial of homeopathy in rheumatoid arthritis. *Rheumatology*. 2001;40(9):1052-5. <http://dx.doi.org/10.1093/rheumatology/40.9.1052>
148. Eldessouki Ibrahim S, el Zohiery A, Mobasher S, Badr Eldin A, Mohamed M, Abdalla A. Acupuncture versus Homeopathy as a Complementary Therapy in Patients with Knee Osteoarthritis. *International Journal of Physical Medicine and Rehabilitation*. 2015;03. 10.4172/2329-9096.1000259
149. Shealy CN, Thomlinson RP, Cox RH, Borgmeyer V. Alternative medicine. Osteoarthritic pain: a comparison of homeopathy and acetaminophen. *American Journal of Pain Management*. 1998;8(3):89-91.
150. Strosser W, Weiser M. Patients with gonarthrosis gaining back mobility - Homoeopathic in a doubleblind comparison. *Biologische Medizin*. 2000;29(6):295-9.
151. van Haselen RA, Fisher PA. A randomized controlled trial comparing topical piroxicam gel with a homeopathic gel in osteoarthritis of the knee. *Rheumatology (Oxford, England)*. 2000;39(7):714-9. 10.1093/rheumatology/39.7.714
152. Widrig R, Suter A, Saller R, Melzer J. Choosing between NSAID and arnica for topical treatment of hand osteoarthritis in a randomised, double-blind study. *Rheumatol Int*. 2007;27(6):585-91. 10.1007/s00296-007-0304-y
153. Khitrov NA. The use of artrofoon in the therapy of disorders of the paraarticular apparatus. *Bull Exp Biol Med*. 2009;148(3):478-81. 10.1007/s10517-010-0741-6
154. Olsen MF, Bjerre E, Hansen MD, Hilden J, Landler NE, Tendal B, et al. Pain relief that matters to patients: systematic review of empirical studies assessing the minimum clinically important difference in acute pain. *BMC Med*. 2017;15(1):35. 10.1186/s12916-016-0775-3
155. Kitchen H, Hansen B, Abetz L, Højbjerg L, Bostrup M. Patient-Reported Outcome Measures For Rheumatoid Arthritis: Minimal Important Differences Review. *Arthritis & Rheumatology*. 2013;65:965.
156. Pope JE, Khanna D, Norrie D, Ouimet JM. The minimally important difference for the health assessment questionnaire in rheumatoid arthritis clinical practice is smaller than in randomized controlled trials. *J Rheumatol*. 2009;36(2):254-9. 10.3899/jrheum.080479
157. Mian AN, Ibrahim F, Scott DL, Galloway J, On behalf of the Tsg. Optimal responses in disease activity scores to treatment in rheumatoid arthritis: Is a DAS28 reduction of >1.2 sufficient? *Arthritis Research & Therapy*. 2016;18(1):142. 10.1186/s13075-016-1028-8
158. Ctri. Effect of Predefined Homoeopathic Medicines in Cervical Spondylosis Pain Management. <https://trialsearchwho.int/Trial2.aspx?TrialID=CTRI/2011/12/002270>. 2011.
159. Morris M, Pellow J, Solomon EM, Tsele-Tebakang T. Physiotherapy and a Homeopathic Complex for Chronic Low-back Pain Due to Osteoarthritis: A Randomized, Controlled Pilot Study. *Altern Ther Health Med*. 2016;22(1):48-56.
160. Stam C, Bonnet MS, Van Haselen RA. The efficacy and safety of a homeopathic gel in the treatment of acute low back pain: A multi-centre, randomised, double-blind comparative clinical trial. *British Homeopathic Journal*. 2001;90(1):21-8.
161. Ostelo RW, de Vet HC. Clinically important outcomes in low back pain. *Best Pract Res Clin Rheumatol*. 2005;19(4):593-607. 10.1016/j.berh.2005.03.003
162. MacDowall A, Skeppholm M, Robinson Y, Olerud C. Validation of the visual analog scale in the cervical spine. *J Neurosurg Spine*. 2018;28(3):227-35. 10.3171/2017.5.Spine1732
163. Johnsen LG, Hellum C, Nygaard OP, Storheim K, Brox JI, Rossvoll I, et al. Comparison of the SF6D, the EQ5D, and the Oswestry disability index in patients with chronic low back pain and degenerative disc disease. *BMC Musculoskelet Disord*. 2013;14:148. 10.1186/1471-2474-14-148

164. Andrade D, Carmona F, Angelucci MA, Martinez EZ, Pereira AMS. Efficacy of a Homeopathic Medicine of *Capsicum frutescens* L. (Solanaceae) in the Treatment of Hot Flashes in Menopausal Women: A Phase-2 Randomized Controlled Trial. *Homeopathy*. 2019;108(2):102-7. 10.1055/s-0038-1676326
165. Colau JC, Vincent S, Marijnen P, Allaert FA. Efficacy of a non-hormonal treatment, BRN-01, on menopausal hot flashes: a multicenter, randomized, double-blind, placebo-controlled trial. *Drugs R D*. 2012;12(3):107-19. 10.2165/11640240-000000000-00000
166. Gupta J, Kulshreshtha D, Lamba CD, Gupta P, Shinde VH, Wadhwa B, et al. Homoeopathic medicine – Sepia for the management of menopausal symptoms: A multicentric, randomised, double-blind placebo-controlled clinical trial. *Indian Journal of Research in Homoeopathy*. 2019;13:219 - 28.
167. Jacobs J, Herman P, Heron K, Olsen S, Vaughters L. Homeopathy for menopausal symptoms in breast cancer survivors: a preliminary randomized controlled trial. *J Altern Complement Med*. 2005;11(1):21-7. 10.1089/acm.2005.11.21
168. von Hagens C, Schiller P, Godbillon B, Osburg J, Klose C, Limprecht R, et al. Treating menopausal symptoms with a complex remedy or placebo: a randomized controlled trial. *Climacteric*. 2012;15(4):358-67. 10.3109/13697137.2011.597895
169. Relton C, O'Cathain A, Nicholl J. A pilot 'cohort multiple randomised controlled trial' of treatment by a homeopath for women with menopausal hot flushes. *Contemp Clin Trials*. 2012;33(5):853-9. 10.1016/j.cct.2012.04.003
170. Constantine GD, Revicki DA, Kagan R, Simon JA, Graham S, Bernick B, et al. Evaluation of clinical meaningfulness of estrogen plus progesterone oral capsule (TX-001HR) on moderate to severe vasomotor symptoms. *Menopause*. 2019;26(5):513-9. 10.1097/gme.0000000000001261
171. Carpenter JS, Bakoyannis G, Otte JL, Chen CX, Rand KL, Woods N, et al. Validity, cut-points, and minimally important differences for two hot flash-related daily interference scales. *Menopause*. 2017;24(8):877-85. 10.1097/gme.0000000000000871
172. Quality Metric Incorporated. User's manual for the SF-36v2 survey. 3rd ed. Lincoln RD, editor 2011.
173. Goldstein KM, Coeytaux RR, Williams JW Jr. Nonpharmacologic Treatments for Menopause-Associated Vasomotor Symptoms [Internet]. Washington (DC): Department of Veterans Affairs (US); 2016.
174. Charandabi Sma Biglu Mh Rad KY. Effect of homeopathy on pain intensity and quality of life of students with primary dysmenorrhea: A randomized controlled trial. *Iranian Red Crescent Medical Journal*. 2016;18(9):Article Number: e30902.
175. Irct201207063706N. Effect of homeopathy on the pain, menstrual symptoms and quality of life in patients with primary dysmenorrhoea. <https://trialsearchwho.int/Trial2.aspx?TrialID=IRCT201207063706N14>. 2012.
176. Ctri. Role of Homoeopathy in Primary Dysmenorrhoea. <https://trialsearchwho.int/Trial2.aspx?TrialID=CTRI/2018/07/014949>. 2018.
177. Teixeira MZ, Podgaec S, Baracat EC. Potentized estrogen in homeopathic treatment of endometriosis-associated pelvic pain: A 24-week, randomized, double-blind, placebo-controlled study. *Eur J Obstet Gynecol Reprod Biol*. 2017;211:48-55. 10.1016/j.ejogrb.2017.01.052
178. Yakir M, Kreitler S, Oberbaum M, Bzizinsky A, Vithoulkas G, Bentwich Z. Homeopathic treatment of premenstrual syndrome: a pilot study. *Unpublished: 8th GIRI Meeting, Jerusalem Israel, December 1994*. 1994:49-50.
179. Yakir M, Kreitler S, Brzezinski A, Vithoulkas G, Oberbaum M, Bentwich Z. Effects of homeopathic treatment in women with premenstrual syndrome: a pilot study. *Br Homeopath J*. 2001;90(3):148-53. 10.1054/homp.1999.0491
180. Yakir M, Klein-Laansma CT, Kreitler S, Brzezinski A, Oberbaum M, Vithoulkas G, et al. A Placebo-Controlled Double-Blind Randomized Trial with Individualized Homeopathic Treatment Using a Symptom Cluster Approach in Women with Premenstrual Syndrome. *Homeopathy*. 2019;108(4):256-69. 10.1055/s-0039-1691834
181. Klein-Laansma CT, Jong M, von Hagens C, Jansen J, van Wietmarschen H, Jong MC. Semi-Individualized Homeopathy Add-On Versus Usual Care Only for Premenstrual Disorders: A Randomized, Controlled Feasibility Study. *J Altern Complement Med*. 2018;24(7):684-93. 10.1089/acm.2017.0388

182. McKendrick M. Chronic fatigue syndrome: a controlled trial of the efficacy of homeopathic treatment. *National Research Register*. 1999.
183. Stanley PJ. Chronic fatigue syndrome: A controlled trial of the efficacy of homeopathic treatment. *National Research Register*. 2001.
184. Weatherley-Jones E, Nicholl JP, Thomas KJ, Parry GJ, McKendrick MW, Green ST, et al. A randomized, controlled, triple-blind trial of the efficacy of homeopathic treatment for chronic fatigue syndrome. *Journal of Psychosomatic Research*. 2004;56(2):189-97. <https://dx.doi.org/10.1016/S0022-3999%2803%2900377-5>
185. Weatherley-Jones E, Thomas K. A randomised, controlled trial of homeopathic treatment for chronic fatigue syndrome. *17th annual meeting of the international society of technology assessment in health care: building bridges between policy, providers, patients and industry; 2001 june 3-6*. 2001:67.
186. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res*. 1995;39(3):315-25. 10.1016/0022-3999(94)00125-o
187. Bell IR, Lewis DA, 2nd, Brooks AJ, Schwartz GE, Lewis SE, Caspi O, et al. Individual differences in response to randomly assigned active individualized homeopathic and placebo treatment in fibromyalgia: implications of a double-blinded optional crossover design. *J Altern Complement Med*. 2004;10(2):269-83. 10.1089/107555304323062266
188. Bell IR, Lewis DA, 2nd, Brooks AJ, Schwartz GE, Lewis SE, Walsh BT, et al. Improved clinical status in fibromyalgia patients treated with individualized homeopathic remedies versus placebo. *Rheumatology (Oxford)*. 2004;43(5):577-82. 10.1093/rheumatology/keh111
189. Bell IR, Lewis DA, 2nd, Schwartz GE, Lewis SE, Caspi O, Scott A, et al. Electroencephalographic cordance patterns distinguish exceptional clinical responders with fibromyalgia to individualized homeopathic medicines. *J Altern Complement Med*. 2004;10(2):285-99. 10.1089/107555304323062275
190. Bell IR, Lewis IDA, Lewis SE, Schwartz GE, Brooks AJ, Scott A, et al. EEG alpha sensitization in individualized homeopathic treatment of fibromyalgia. *International Journal of Neuroscience*. 2004;114(9):1195-220. <https://dx.doi.org/10.1080/00207450490475724>
191. Fisher P. Rhus toxicodendron in the treatment of fibromyalgia: a double-blind, placebo-controlled trial, with cross-over. *J Omhi*. 1988;1(3):26-8.
192. Fisher P, Greenwood A, Huskisson EC, Turner P, Belon P. Effect of homeopathic treatment on fibrositis (primary fibromyalgia). *BMJ (Clinical research ed)*. 1989;299(6695):365-6.
193. Relton C, Smith C, Raw J, Walters C, Adebajo AO, Thomas KJ, et al. Healthcare provided by a homeopath as an adjunct to usual care for Fibromyalgia (FMS): results of a pilot Randomised Controlled Trial. *Homeopathy*. 2009;98(2):77-82. 10.1016/j.homp.2008.12.004
194. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain*. 1975;1(3):277-99. 10.1016/0304-3959(75)90044-5
195. Lin S, Hsiao Y-Y, Wang M. Test Review: The Profile of Mood States 2nd Edition. *Journal of Psychoeducational Assessment*. 2014;32(3):273-7. 10.1177/0734282913505995
196. Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. *Arthritis Res Ther*. 2009;11(4):R120. 10.1186/ar2783
197. Bennett RM, Bushmakina AG, Cappelleri JC, Zlateva G, Sadosky AB. Minimal clinically important difference in the fibromyalgia impact questionnaire. *J Rheumatol*. 2009;36(6):1304-11. 10.3899/jrheum.081090
198. Rishi P, Rishi E, Maitray A, Agarwal A, Nair S, Gopalakrishnan S. Hospital anxiety and depression scale assessment of 100 patients before and after using low vision care: A prospective study in a tertiary eye-care setting. *Indian J Ophthalmol*. 2017;65(11):1203-8. 10.4103/ijo.IJO_436_17
199. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. 10.1136/bmj.l4898
200. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors.

Cochrane Handbook for Systematic Reviews of Interventions version 6.0. (updated July 2019): Cochrane 2019.

201. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002. 10.1136/bmj.d4002