Systematic Review and Meta-Analysis of Randomised, Other-than-Placebo Controlled, Trials of Non-Individualised Homeopathic Treatment

Robert T. Mathie[1] Yvonne Y.Y. Fok[2] Petter Viksveen[3] Aaron K.L. To[2] Jonathan R.T. Davidson[4]

[1] Homeopathy Research Institute, London, United Kingdom
[2] The Hong Kong Association of Homeopathy, Hong Kong
[3] Faculty of Health Sciences, University of Stavanger, Stavanger, Norway
[4] Department of Psychiatry and Behavioral Science, Duke University Medical Center, Durham, North Carolina, United States

Address for correspondence Robert T. Mathie, PhD, Homeopathy Research Institute, London, United Kingdom (e-mail: robertmathie@hri-research.org).

Homeopathy 2019;108:88–101.

Abstract

Introduction This study focuses on randomised controlled trials (RCTs) of non-individualised homeopathic treatment (NIHT) in which the control (comparator) group was other than placebo (OTP).

Objectives To determine the comparative effectiveness of NIHT on health-related outcomes in adults and children for any given condition that has been the subject of at least one OTP-controlled trial. For each study, to assess its risk of bias and to determine whether its study attitude was predominantly 'pragmatic' or 'explanatory'.

Methods Systematic review. For each eligible trial, published in the peer-reviewed literature up to the end of 2016, we assessed its risk of bias (internal validity) using the seven-domain Cochrane tool, and its relative pragmatic or explanatory attitude (external validity) using the 10-domain *PRECIS* tool. We grouped RCTs by whether these examined IHT as alternative treatment (study design 1a), adjunctively with another intervention (design 1b), or compared with no intervention (design 2). RCTs were sub-categorised as superiority trials or equivalence/non-inferiority trials. For each RCT, we designated a single 'main outcome measure' to use in meta-analysis: 'effect size' was reported as odds ratio (OR; values > 1 favouring homeopathy) or standardised mean difference (SMD; values < 0 favouring homeopathy).

Results Seventeen RCTs, representing 15 different medical conditions, were eligible for study. Three of the trials were more pragmatic than explanatory, two were more explanatory than pragmatic, and 12 were equally pragmatic and explanatory. Fourteen trials were rated 'high risk of bias' overall; the other three trials were rated 'uncertain risk of bias' overall. Ten trials had data that were extractable for analysis. Significant heterogeneity undermined the planned meta-analyses or their meaningful interpretation. For the three equivalence or non-inferiority trials with extractable data, the small, non-significant, pooled effect size (SMD $= 0.08$; $p = 0.46$) was consistent with a conclusion that NIHT did not differ from treatment by a comparator (*Ginkgo biloba* or betahistine) for vertigo or (cromolyn sodium) for seasonal allergic rhinitis.

Keywords
- comparative effectiveness
- explanatory trial
- non-individualised homeopathic treatment
- meta-analysis
- pragmatic trial
- randomised controlled trial
- risk of bias
- systematic review

received
August 1, 2018
accepted after revision
November 9, 2018
published online
January 30, 2019

DOI https://doi.org/10.1055/s-0038-1677481.
ISSN 1475-4916.

Conclusions The current data preclude a decisive conclusion about the comparative effectiveness of NIHT. Generalisability of findings is restricted by the limited external validity identified overall. The highest intrinsic quality was observed in the equivalence and non-inferiority trials of NIHT.

Introduction

Homeopathy is a system of medicine that uses specific preparations of substances whose effects, when administered to healthy subjects, correspond to the manifestations of the disorder (symptoms, clinical signs, pathological states) in the individual patient.[1] In *individualised homeopathy*—as originally defined by its founder, Samuel Hahnemann—typically a single homeopathic medicine is selected on the basis of the 'symptom picture' of a patient. In clinical homeopathy, by contrast, one or more homeopathic medicines are administered for standard clinical situations or conventional diagnoses. In complex homeopathy, several homeopathic medicines are combined in a fixed ('complex') formulation. Isopathy is the use of homeopathic dilutions from the causative agent of the disease itself, or from a product of the disease process, to treat the condition.[1] In the context of a randomised controlled trial (RCT), none of the latter three approaches involves matching a patient with the 'total symptom picture' of an individually prescribed homeopathic medicine: each is thus termed *non-individualised homeopathy*.

The current systematic review (SR) focuses on RCTs of non-individualised homeopathic treatment (NIHT), in which the control (comparator) group was something other than placebo (OTP). Two essentially different options exist for OTP study design of RCTs: (1) other therapeutic intervention (e.g. a conventional medicine or a physical therapy), which can be sub-divided into (a) trials in which NIHT is given as an *alternative* to the comparator intervention, and (b) trials in which NIHT *combined with the other intervention* is compared with the other intervention alone (the '[A + B] versus B' approach); (2) no therapeutic intervention (usually waiting-list controls). Trials of type 1 can be regarded as 'comparative effectiveness' studies. We have previously conducted an SR of OTP-controlled RCTs of *individualised* homeopathy,[2] as well as two SRs of *placebo-controlled* RCTs of homeopathy.[3,4]

No previous SR has considered solely RCTs of non-individualised homeopathy that were controlled by an OTP intervention. One SR of OTP-controlled trials did focus on *individualised* homeopathy:[5] published in 1999, it identified six eligible trials, two of which favoured homeopathy, two favoured conventional drugs, and two were non-conclusive. That review concluded overall that the 'value of individualized homoeopathy relative to allopathic treatments' was not known.

An OTP-controlled trial would usually be intended to possess a 'pragmatic' rather than an 'explanatory' study attitude.[6] A pragmatic trial tends to promote *external validity* (generalisability of results to a relevant population of patients).[7,8] Though it is seldom formally approached by trialists or in SRs,[6,9] external validity can be addressed in a tool (*PRECIS*), developed by Thorpe et al in 2009, that assesses a trial design's positioning on 'the pragmatic—explanatory continuum'.[10] For the current SR of OTP-controlled trials of NIHT, we assess each study's internal validity as well as its external validity, with a view to appraising both its intrinsic quality and the extent of its pragmatic/explanatory attitude.

Aim of the Study

Our objective was to examine comparative effectiveness of NIHT in OTP-controlled trials of any clinical condition, in adults or children, for which there was at least one eligible RCT. Using meta-analysis, we aimed to evaluate RCTs that have investigated NIHT: (study design 1a) in comparison to another therapeutic intervention; (study design 1b) adjunctively with another treatment intervention, in comparison to that other intervention alone ('[A + B] versus B'); or (study design 2) compared with no other intervention. RCTs were further sub-categorised as superiority trials or equivalence/non-inferiority trials. An additional aim was, if possible, to evaluate by meta-analysis the comparative effectiveness of NIHT for any given clinical condition or category of conditions. In all cases, we reflected matters of internal validity (risk of bias) and external validity (pragmatic/explanatory study attitude).

Methods

Methods complied fully with the *PRISMA* 2009 checklist (►**Supplementary File 1** [available in online version only]) and with our published study protocol.[11]

Search Strategy, Data Sources and Trial Eligibility

We conducted a systematic literature search to identify RCTs that compared NIHT with something other than placebo, for any clinical condition.[12] Each of the following electronic databases was searched from its inception up to the end of 2011, with updated searches of the same databases up to the end of 2016: AMED; CAM-Quest; CINAHL; Cochrane Central Register of Controlled Trials; Embase; Hom-Inform; LILACS; PubMed; Science Citation Index; Scopus. For the update, CORE-Hom was also searched, using the term 'randomised' or 'unknown' in the *Sequence Generation* field.

The full electronic search strategy for PubMed (Cochrane Highly Sensitive Search Strategy) was conducted as per the above: '((homeopath* *or* homoeopath*) *and* ((randomised controlled trial[pt]) *or* (controlled clinical trial[pt]) *or* (randomised[tiab]) *or* (placebo [tiab]) *or* (clinical trials as topic [mesh:noexp]) *or* (randomly[tiab]) *or* (trial[ti]))) *not* (animals[mh] *not* humans[mh])'.

Specific pre-defined exclusion criteria were then applied:

- Trials of homeopathic prophylaxis
- Trials with crossover design
- Research using radionically prepared 'homeopathic' medicines[13]
- The tested intervention is NIHT in tandem with other (complementary or conventional) medicine or therapy, and where the nature of the combined comparator intervention makes it impossible to distinguish any effects due to NIHT[a]
- Other specified reason.

Whereas a placebo-controlled trial of non-individualised homeopathy can be fully blinded, it is more difficult—and sometimes impossible—to achieve such blinding in a corresponding OTP-controlled trial. Unlike the case for our corresponding SR of placebo-controlled trials,[4] therefore, patient- and/or practitioner-unblinded trials *are* eligible for the current SR of OTP trials.

Only published data were eligible for analysis. The authors of eligible articles were not approached for clarification on unclear or missing facts;[14] however, original authors' cross-reference to their previously published study methods were followed up and taken into account as necessary. For trials with more than two study groups, and where such trials had not previously been catalogued under 'placebo-controlled', only the data concerning comparisons between NIHT and OTP were extracted from the articles; in relevant cases of more than one OTP control, a study group comprising actual treatment was favoured for analysis over one comprising 'no treatment'.

Outcome Definitions

For each trial, and for the purposes of risk-of-bias assessment, we identified a 'main outcome measure' using a refinement of the approaches adopted by Linde et al and by Shang et al.[15,16] As for our previous SRs,[2–4] and per protocol,[11] each trial's 'main outcome measure' was identified based on a hierarchical ranking order (consistent with the WHO ICF Classification System for Levels of Functioning Linked to Health Condition[b]).

We followed the WHO ICF system without reference to any 'primary outcome measure' that might have been identified by the original investigators. Unless otherwise indicated, the single end-point (measured from the start of the intervention) associated with our designated 'main outcome measure' was taken as the last follow-up at which data were reported for that outcome.

[a]This study design is distinct from the eligible '[A + B] versus B' design, and from eligible studies that allow concomitant conventional medication to remain ongoing in the subjects of *each* study group.

[b]Towards a Common Language for Functioning, Disability and Health. ICF: The International Classification of Functioning, Disability and Health. Geneva; World Health Organization, 2002.

Data Extraction

Two assessors (RTM and YYYF, or RTM and PV) independently extracted relevant data using a standard data recording approach, in spreadsheet format (Microsoft *Excel*). The data extracted per trial included, as appropriate: demographics of participants (gender, age range, clinical condition); study setting; potency or potencies of homeopathic medicines; dosage frequency; whether a pilot trial; 'main outcome measure' (see below) and measured end-point; other outcome measures reported; funding source/s. The statistical items noted were: sample size and missing data for each intervention group; whether power calculation carried out; whether intention-to-treat (ITT), per-protocol, complier-average-causal-effect,[17] or other type of primary analysis.

Assessment of Risk of Bias (Internal Validity)

Using the standard criteria defined by Cochrane,[14] the extraction of information enabled appraisal of 'low risk', 'uncertain risk' or 'high risk' of bias with respect to: (domain I) the methods used to generate the random sequence; (domain II) the method of allocation concealment used to implement the random sequence; (domain IIIa) the blinding of participants and/or study personnel; (domain IIIb) the blinding of outcome assessors; (domain IV) completeness of the outcome data included in the analysis; (domain V) evidence of selective outcome reporting; (domain VI) evidence of other bias, including data imbalance between the groups at baseline.

Two or three assessors (RTM and PV, or RTM and YYYF and AKLT) carried out their assessments independently, with discrepancies between them resolved by consensus discussion and, if necessary, the input of another co-author. For domain IV, a trial was normally regarded as no better than 'unclear' if there was greater than 20% participant attrition rate, irrespective of whether ITT analysis had been carried out. Domain V was automatically attributed 'high risk of bias' if its designated main outcome measure could not be extracted to enable calculation of 'relative effect size' (see below). The nature of any research sponsorship was taken into account for sub-group analysis (see below), not in risk-of-bias assessment *per se*.

Rating of Trials for Risk of Bias

As per the standard Cochrane approach, each trial was designated: *low risk of bias* for all key domains; or *uncertain risk of bias* for one or more key domains; or *high risk of bias* for one or more key domains.[14] We used our novel method of nomenclature, based on the Cochrane approach, for rating risk-of-bias characteristics across all domains per trial:[2–4]

A = *Low risk of bias* in all seven domains.

B*x* = *Uncertain risk of bias* in *x* domains; low risk of bias in all other domains.

C*y.x* = *High risk of bias* in *y* domains; uncertain risk of bias in *x* domains; low risk of bias in all other domains.

An 'A'-rated trial was designated *reliable evidence*. We also designated a 'B1'-rated trial *reliable evidence* (and listed as 'B1* [minimal risk of bias]') if the uncertainty in its risk

of bias was for one of domains IV, V or VI only (i.e. it is required to be judged free of bias for each of domains I, II, IIIa and IIIb).[11]

It is expected that an OTP-controlled trial would be rated 'high risk of bias' in assessment domain IIIa. We recognise that this is a normal feature of an OTP-controlled trial, and which thus inevitably limits its internal validity—see also the section *Sensitivity Analyses*, below.

Assessment of Pragmatic/Explanatory Attitude (External Validity)

Equating external validity to study attitude, we adopted the *PRECIS* approach[10] to assess each trial's positioning on the pragmatic—explanatory continuum, taking account of 10 domains:

1. Participant eligibility criteria;
2. Experimental intervention flexibility;
3. Practitioner expertise—experimental intervention;
4. Comparison intervention;
5. Practitioner expertise—comparison intervention;
6. Follow-up intensity;
7. Primary trial outcome;
8. Participant compliance with 'prescribed' intervention;
9. Practitioner adherence to study protocol;
10. Analysis of primary ('main') outcome.

Against a set of standard judgmental criteria,[10] two of us (RTM and YYYF, or RTM and PV) independently assessed each of the 10 attributes as 'Much more explanatory than pragmatic', 'More explanatory than pragmatic', 'Equally pragmatic and explanatory', 'More pragmatic than explanatory' or 'Much more pragmatic than explanatory'—see details in our previous SR article.[2] Consistent with the *PRECIS* authors' methodological approach, we devised our own notation, giving a score of 1, 2, 3, 4 or 5 in each of the above five cases respectively.[2] Where an article contained no relevant information to enable assessment of a given domain, a score of 3 was given. The range of possible total scores per trial was 10 to 50.

We characterised each trial overall using the following empirical consideration of its total score:

10–17: Much more explanatory than pragmatic;
18–25: More explanatory than pragmatic;
26–34: Equally pragmatic and explanatory;
35–42: More pragmatic than explanatory;
43–50: Much more pragmatic than explanatory.

Discrepancies between assessments (per attribute and by overall characterisation per trial) were resolved by consensus discussion. Initial attention focused on each domain for which the two independent scores were a value of 2 or more apart, and with the aim of narrowing the discrepancy to no more than a value of 1. Subsequent discussion aimed to re-appraise any studies whose two sets of total scores lay on either side of a threshold for overall characterisation of pragmatic/explanatory attitude, with a view to agreeing a final designation; we did not aim to calculate the mean value of the two independent total *PRECIS* scores.

Study Selection for Meta-Analysis

All RCTs that were included in the SR were potentially eligible for meta-analysis. If the original article did not provide or inform adequate data on the selected 'main outcome measure' to enable extraction or calculation of the standardised mean difference (SMD) or the odds ratio (OR), we described the selected main outcome as 'not estimable': an alternative, estimable, outcome was not sought.

Consistent with the above, the following studies were excluded from meta-analysis:

- Those that presented non-parametric data only, and where there was no information that enabled the data distribution to be assessed;
- Those from which the necessary data could not be extracted (not provided or uninterpretable).

Summary Measures for 'Main Outcome'

For the remaining relevant records of NIHT, we aimed to examine (1) overall relative effect sizes; (2) relative effect sizes by disease; (3) relative effect sizes by disease category. In each of these three cases, 'relative effect size' was taken as the difference (if relevant—see below) between the homeopathy and the control groups at the identified end-point of the trial, and using per-protocol data:[4]

- For dichotomous measures: OR, with 95% confidence interval (CI);[c]
- For continuous measures: SMD, with 95% CI.[d]

In trials where the main outcome measure was a continuous variable, and where there were insufficient data presented to identify the mean and/or the standard deviation (SD) per group at the defined end-point, the necessary data were estimated, if possible, by imputing relevant other data from the same study.[18]

Statistical Interpretation per Study Design

The interpretation of a statistical finding of $p < 0.05$ (direction of effect towards homeopathy or towards control) and of $p > 0.05$ (direction of effect towards either homeopathy or control) has been detailed in our recent SR of OTP-controlled trials of individualised homeopathic treatment.[2] We separately interpreted the findings from 'equivalence' or 'non-inferiority' trials,[19,20] reflecting the original authors' margin of equivalence or non-inferiority as appropriate.

For any RCT or group of RCTs on a given clinical condition/category, the interpretation of NIHT as 'effective', 'ineffective' or 'inconclusive' applied solely to the particular clinical condition/category examined.

Synthesis of Quantitative Results

Overall 'Relative Effect Size' of NIHT

For groups of eligible RCTs that have compared NIHT (1a) with another intervention, or (1b) adjunctively with another intervention, or (2) with no treatment, we aimed to pool for

[c] OR > 1 favours homeopathy.
[d] SMD < 0 favours homeopathy.

meta-analysis the 'main outcome' data in two separate sets of studies as appropriate, using either the OR or the SMD of each relevant trial.[21] For each study design (1a, 1b, 2), it was then planned to combine data—if sufficient in number—from the two sets of studies (OR and SMD) into a single forest plot, re-expressing SMDs by transformation to OR, using an approximation method proposed by Chinn[22] and recommended by the Cochrane Statistical Methods Group.[23]

Based on the assumption of at least moderate clinical heterogeneity among studies, the 'random effects' statistical model for meta-analysis was selected rather than the 'fixed effect' model.[21]

Disease- and Category-Specific 'Relative Effect Size' of NIHT

For each specific clinical condition or category of conditions, for each of study designs 1a, 1b and 2, and for which there was >1 RCT of given type and with extractable main outcome, we planned to pool the data using meta-analysis.

Heterogeneity and Asymmetry

The I^2 statistic was used to assess the variability between studies in a given meta-analysis: it gives the percentage of the total variability in the estimated effect size (which is composed of between-study heterogeneity plus sampling variability) that is attributable to heterogeneity. The I^2 statistic can take values between 0% and 100%: $I^2 = 0\%$ means that all of the heterogeneity is due to sampling error; $I^2 = 100\%$ means that all variability is due to true heterogeneity between studies. Where feasible, it was intended to use funnel plots to assess the impact of publication bias.

Additional Quantitative Analyses on Overall 'Relative Effect Size' of NIHT

Sensitivity Analyses

We planned sensitivity analyses based separately on (1) our risk-of-bias ratings and on (2) our assessments of external validity. We aimed to reflect in this analysis any trial that was categorised as '(much) more pragmatic than explanatory' and whose internal validity was compromised by high risk of bias *in domain IIIa only* (blinding of participants and/or study personnel).

Sub-Group Analyses

Comparative forest plots were planned—for each of study designs 1a, 1b and 2—on the following sub-groups of trial attributes:

- Whether or not a pilot (or 'preliminary' or 'feasibility') study, as defined by the original authors;
- Whether or not sample size > median for all trials with extractable data;
- Whether or not potency/potencies of homeopathic medicines ≥12C;
- Whether or not the research sponsor is an organisation (e.g. homeopathic pharmacy) with potential vested interest in the trial findings.

Results

Included Studies

The *PRISMA* flowchart from the original comprehensive literature search (up to and including 2011) was published previously.[12] An updated *PRISMA* flowchart is given in ►**Fig. 1**, identifying a total of 588 records;[e] 488 remained after removal of duplicates. After excluding 102 due to type of record (book chapters, theses, abstracts and other minor articles), 386 full-text records were then assessed for eligibility. Three hundred and sixty-nine records were excluded for the general reasons summarised in ►**Fig. 1**; nine of these same 369 were excluded from the present SR for the additionally specified reasons shown in ►**Supplementary File 2** (available in online version only) (e.g. mother tincture in homeopathic formulation). The finally remaining 17 records (17 RCTs) were thus included in this SR; data were not extractable from 7 of those (►**Supplementary File 2** [available in online version only]),[f] leaving 10 records available for quantitative analysis.

Characteristics of Included Studies

The 17 RCTs represented 15 clinical conditions across 13 categories (►**Table 1**). Homeopathic potency was ≥12C in three of the 15 trials for which relevant information was available. Five trials were free of vested interest; five trials were not free of vested interest; seven trials did not enable certainty in this assessment.

Summary of Findings

For each trial, ►**Table 2** includes details of the sample size, the identified main outcome measure (and whether dichotomous or continuous), the end-point, and a number of other study attributes. Two trials were described in the original article as a 'pilot' (or 'preliminary' or 'feasibility') study. A power calculation was carried out for eight of the trials. ITT was the basis for the original analysis in four trials. Mean attrition rate per RCT was 7.4%. The main outcome variable was dichotomous in four studies and continuous in the other 13. Only 6 of the 17 original RCTs clearly described a 'main' or 'primary' outcome (A153, Karow; A159, Taylor; A163, Weiser; A297, Mourão; A304, Thinesse-Mallwitz; A319, Jong); in each of the 6 cases, this corresponded to our designated 'main outcome measure'. The total sample size was 1,376 for the 10 trials whose data were amenable to quantitative analysis; their median sample size was 102 (inter-quartile range: 56.5 to 144.5). The 10 analysable studies included 10 different main outcome measures and for an end-point that ranged from 4 days to 2 years.

Risk of Bias

►**Table 3** provides the risk-of-bias details for each of the 17 trials, and sub-divided by study design: (1a) the 13 with other-intervention control; (1b) the four that were '[A + B] versus B'.

[e] Complete details of all 588 records are available from the website of the Homeopathy Research Institute.

[f] ►**Supplementary file 2** (available in online version only) represents an update of the flowchart that is included in the study protocol.[11]

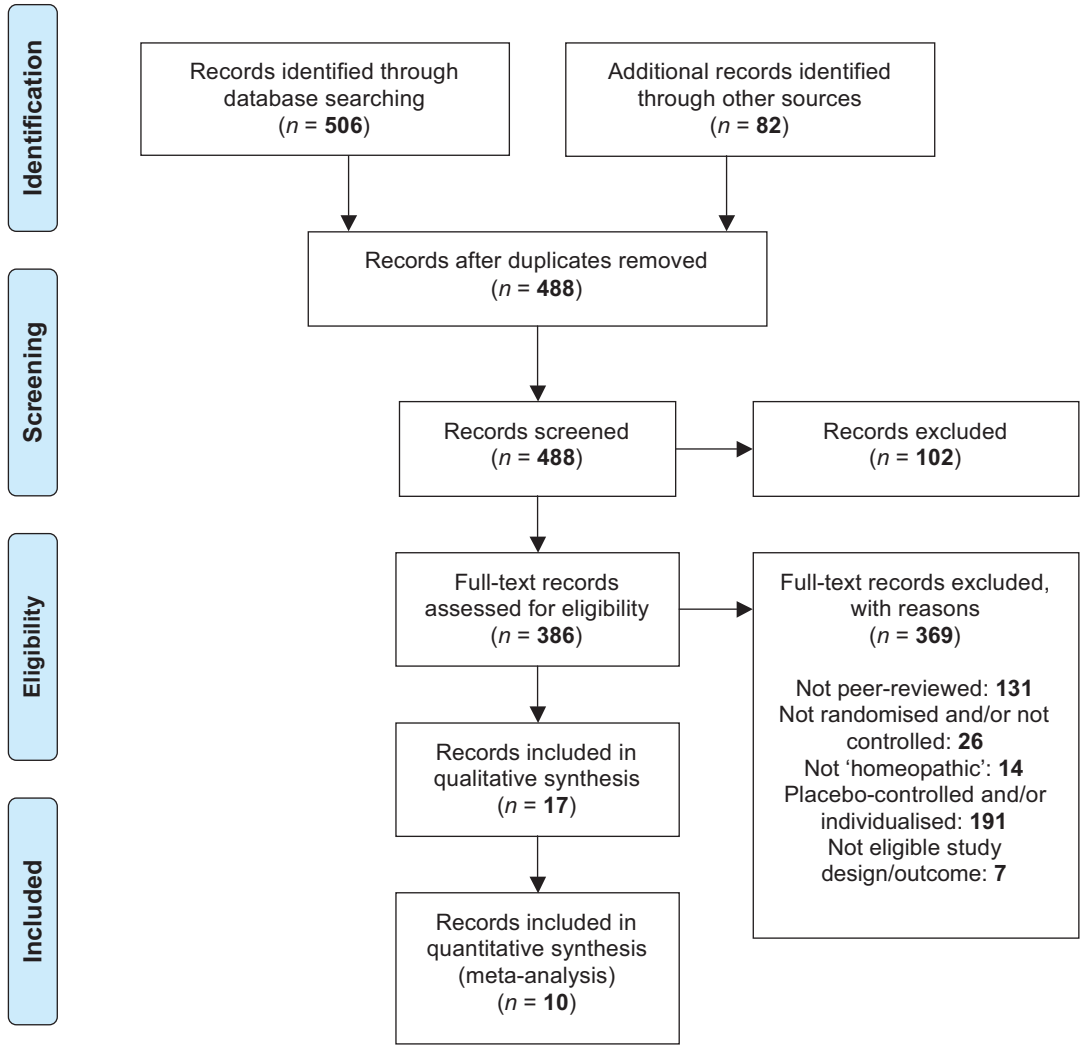

Fig. 1 Updated *PRISMA* flowchart for all records published up to and including 2016.

Domain IIIa (blinding of participants and/or study personnel), domain IIIb (blinding of outcome assessors) and domain V (selective outcome reporting) presented the greatest methodological concerns. The seven trials with *high risk of bias* for domain V were so because they did not contain data that were extractable for meta-analysis. Domain I (sequence generation) and domain II (allocation concealment) presented the most uncertain methodological judgments.

There were no A- or B1*-rated trials. Three trials (A151, Issing; A162, Weiser; A163, Weiser) were rated *uncertain risk of bias* ('B1'- or 'B2'-rated). Fourteen trials were assessed as *high risk of bias* ('C'-rated); one of those (A297, Mourão) was judged *high risk of bias* for domain IIIa only. Seven 'C'-rated trials were deficient in two or more domains of assessment ('C2.0'-rated or worse). A summary risk-of-bias bar-graph is shown in ►**Supplementary File 3** (available in online version only).

Pragmatic/Explanatory Attitude

Independent assessment identified a total of 52 domains (2 to 6 domains per trial) for which the two scores were ≥ 2 apart. In all cases, consensus discussion narrowed the score discrepancy per domain to no more than 1, which was our target. There were two trials for which the consequent two total scores lay on either side of a threshold for overall characterisation of pragmatic/explanatory attitude: however, a final designation per trial was readily achieved through further consensus discussion (►**Table 3**).

Three of the trials were more pragmatic than explanatory, 2 were more explanatory than pragmatic, and 12 were equally pragmatic and explanatory. One of the three trials categorised as 'more pragmatic than explanatory' (A287, Villanueva) had high risk of bias in domain IIIa only (blinding of participants and/or study personnel).

Across all trials, each of 47 domains was given a score of 3 by both assessors due to the absence of sufficient information in the original article: this lack of information pertained especially to domains 3 (practitioner expertise—homeopathy), 5 (practitioner expertise—comparison intervention), 8 (participant compliance) and 9 (protocol adherence).

Meta-Analysis

The selected main outcome was 'not estimable' for seven studies (►**Supplementary File 2** [available in online version only]), for which group SDs for the designated end-point could

Table 1 Demographic data for 17 OTP-controlled RCTs of non-individualised homeopathic treatment: seven studies excluded from meta-analysis shown in italics

No.	First author	Year	Study design	Category	Condition	Participants' demographics	Study setting	Hom. med., potency	Hom. dosage freq.	Comparator	Funding source	Free of vested interest[a]
A150	Gassinger	1981	1a	Respiratory infection	Common cold	Aged 20–70 yrs	Medical clinic	*Eupatorium perfoliatum* D2	5 × 10 drops or 5 × 2 tablets	Acetylsalicylic acid: 3 × 1 tablet	None stated	U
A151	Issing	2005	1a	Nausea/ Vertigo	Vertigo	M or F patients, 60–80 yrs, with at least 3 episodes of vertigo/ day in week prior to study	13 study centres (clinics practicing CAM and/or conventional medicine) in Germany	*Cocculus* D4, *Conium* D3, *Ambra* D6, *Petroleum* D8	Two tablets t.i.d. for 8 wk	*Ginkgo biloba* 24%: one tablet t.i.d. for 8 wk	Biologische Heilmittel Heel GmbH, Germany	N
A153	*Karow*	*2008*	1a	Surgery & anaesthesiology	Post-operative wound healing	M or F, 20–65 yrs, with surgical indication 'Hallux valgus' on left and/or right big toe	Orthopaedic hospital, Germany	*Arnica* D4	10 pillules, 3 times per day, for 4 d post-op	Diclofenac sodium, 50 mg, 3 times daily, for 4 d post-op.	None stated	U
A155	Maiwald	1988	1a	Respiratory infection	Common cold	Army soldiers, aged 17–49	Single centre-clinic, Germany	*Aconitum* D4, *Bryonia* D4, *Lachesis* D12, *Eupatorium perfoliatum* D3, *Phosphorus* D5	3 × 3 tablets, 10 days	Acetylsalicylic acid	None stated	U
A156	*Shealy*	*1998*	1a	Rheumatology	Osteoarthritis	M or F, 34–85 yrs, with confirmed osteoarthritis of one or more joints	Institute of 'comprehensive healthcare and clinical research', USA	*Rhus tox* 12X, *Causticum* 12X, *Lac vac* 30X (liquid)	10 drops, 4 times per day, for 1 mo	Acetaminophen, 2,600 mg daily	Homeopathic preparation 'supplied by Beaumont Laboratories'	U
A159	Taylor	2011	1b	Ear, nose & throat	Otitis media (acute)	Children, 6–11 yrs, diagnosed with acute otitis media	University paediatric medical centre, USA	*Pulsatilla*, *Chamomilla*, *Sulphur*, *Calc carb*, *Belladonna*, *Lycopodium*; all 30c	3–4 drops, up to 3 times/day, as needed for up to 5 d	Standard therapy: antibiotic, plus acetaminophen, ibuprofen, or benzocaine ear drops	Standard Homeopathic Company, manufacturers of the test agent: had no role in data collection, analysis or interpretation	Y
A160	*Totonchi*	*2007*	1a	Surgery & anaesthesiology	Post-operative ecchymosis	M or F, 15–65 yrs, who had had primary rhinoplasty with osteotomy	University Department of Plastic Surgery, USA	*Arnica*, potency not stated	3 times a day for 4 d	Dexamethasone i. v. intra-op, then methyl-prednisone p.o. for 6 d	None stated	U
A162	Weiser	1998	1a	Nausea/ Vertigo	Vertigo	M or F, with acute or chronic vertigo	17 general practices in Germany	*Ambra grisea*, *Cocculus*, *Conium*, *Petroleum*; D3 to D8	15 drops, 3 times daily, for 42 d	Betahistine, 18 mg per day	Study conducted by manufacturer of test agent	N
A163	Weiser	1999	1a	Allergy & asthma	Seasonal allergic rhinitis	M or F outpatients, 18–60 yrs, diagnosed with seasonal allergic rhinitis	17 'study centers' in Germany: 10 GPs; 5 ENT specialists; 2 private practice internists	Nasal Spray: *Luffa operculata*, *Galphimia glauca*, histamine, *Sulphur*; 4X to 200X	One spray, 4 times daily, for 6 wk	Cromolyn sodium, 20 mg/mL	Heel GmbH, manufacturer of the test agent	N
A164	*Wiesenauer*	*1987*	1a	Cardiovascular	Orthostatic hypotension	M or F, chronic hypotension (>3 mo)	15 physicians with or without specialisation in homeopathy	*Haplopappus* D2	3 times (daily?)	Etilefrin	Robert-Bosch-Stiftung; Stiftung zur Förderung der Erfahungsheilkunde	Y

Table 1 (*Continued*)

No.	First author	Year	Study design	Category	Condition	Participants' demographics	Study setting	Hom. med., potency	Hom. dosage freq.	Comparator	Funding source	Free of vested interest[a]
A285	*González de Vega*	*2013*	1a	Musculoskeletal	Sprain	M or F, 18-40 yrs, with moderate to severe acute ankle sprain within the past 24 h	15 out-patient centres in Spain	12 homeopathic medicines, D2 to D6	2 g, 3 times daily	1% diclofenac sodium gel	Sponsored by Biologische Heilmittel Heel GmbH, manufacturer of the test agents	N
A286	*Pellow*	*2013*	1a	Dermatology	Nappy rash	Babies, aged 3-24 mo, with nappy rash	University training centre for homeopathy, South Africa	*Belladonna* 6cH, *Calendula officinalis* D1, *Sulphuricum acidum* 6cH	During nappy changing and after every bath	Milking cream	University of Johannesburg	Y
A287	Villanueva	2012	1b	Miscellaneous	Malnutrition	Malnourished children aged 1-19 yrs	Polyclinic, Cuba	*Calcarea fluorica*, *Calcarea carbonica*, *Calcarea phosphorica*; all 30cH	Not stated	Diet enhancement and poly-vitamins	None stated	U
A297	Mourão	2014	1b	Oral/dental	Chronic periodontitis	M or F, 35-70 yrs, with chronic periodontitis	University dental clinic, Brazil	*Berberis* 6cH, *Mercurius solubilis* 6cH, *Belladona*, *Hepar sulph* 6cH, *Pyrogenium* 200cH	Various	Conventional periodontal therapy: personal oral hygiene, supragingival scaling	'The study was self-funded by the authors and their institutions'	Y
A304	Thinesse-Mallwitz	2015	1b	Respiratory infection	URTI	M or F, aged 1–65 yrs, with URTI diagnosis and fever >37.5C	12 centres in Ukraine (4 GPs, 8 pediatricians); 10 centres in Germany (4 GPs, 6 pediatricians)	*Aconitum* D3, *Bryonia* D2, *Eupatorium perfoliatum* D1, *Gelsemium* D3, *Ipecacuanha* D3, *Phosphorus* D5	Various	Standard treatment (paracetamol, ambroxol, and/or oxymetazoline)	Co-authors received fee from Deutsche Homöopathie-Union, manufacturers of test agent	U
A305	Chakraborty	2015	1a	Tropical disease	Leprosy	Patients aged 14–60 yrs, with confirmed borderline tuberculoid leprosy	Clinical research institute, India	*Sulphur* 200c	Once per week for 2 yrs	Standard treatment: Rifampicin 600 mg once monthly, dapsone 100 mg daily for 6 mo	Ministry of AYUSH, Government of India	Y
A319	*Jong*	*2016*	1a	Mental disorder	Insomnia	M or F children, aged up to 6 yrs, with difficulties falling asleep and maintaining sleep	Five out-patient paediatric clinics in Russia	*Cypripedium pubescens* D4, *Magnesium carbonicum* D10, *Zincum valerianicum* D12	One tablet, four times a day, for 4 wk	Glycine: one tablet, twice a day	Deutsche Homöopathie-Union, manufacturers of the test agent	N

Abbreviations: ENT, ear nose and throat; N, no; OTP, other-than-placebo; RCTs, randomised controlled trials; URTI, upper respiratory tract infection; U, unclear; Y, yes.
[a]Vested interest: Support (direct, through research sponsorship; indirect, via gifted medicines) from company that provided homeopathic medicines for the trial.

Table 2 Summary of findings table: seven studies excluded from meta-analysis shown in italics

No.	First author	Year	Pilot study	Power calc.	ITT sample	PP sample	Attrition rate %	Original primary analysis	'Main' outcome identified	Nature of 'main' outcome	End-point
Study design 1a: NIHT versus other intervention											
A150	Gassinger	1981	N	N	53	53	0.0	PP	Sum of symptoms	Continuous	10 d
A151	[a] Issing	2005	N	N	170	154	9.4	PP	Frequency of vertigo episodes (per day)	Continuous	6 wk
A153	[b] *Karow*	*2008*	N	Y	88	88	0.0	PP	Post-op irritation (VAS, 0-100)	Continuous	4 d
A155	Maiwald	1988	N	Y	170	115	32.4	PP	Proportion with symptom score improvement	Dichotomous	4 d
A156	*Shealy*	*1998*	N	N	65	65	0.0	PP	Average pain (VAS, 0-100)	Continuous	30 d
A160	*Totonchi*	*2007*	N	N	32	32	0.0	PP	Extent of ecchymosis (scale of 0-5)	Continuous	8 d
A162	[b] **Weiser**	1998	N	Y	119	105	11.8	PP	Frequency of vertigo episodes (per day)	Continuous	6 wk
A163	[b] **Weiser**	1999	N	Y	146	135	7.5	ITT	RQLQ	Continuous	6 wk
A164	[b] ***Wiesenauer***	*1987*	N	N	50	41	18.0	ITT	Summary score of subjective symptoms	Continuous	6 wk
A285	[a] ***González de Vega***	*2013*	N	Y	449	385	14.3	ITT	Ankle pain (VAS 0-100)	Continuous	6 wk
A286	*Pellow*	*2013*	Y	N	40	37	7.5	PP	Total percentage area affected (according to the Modified Lund and Browder Chart)	Continuous	7 d
A305	Chakraborty	2015	N	N	60	60	0.0	PP	Presence of skin sensation	Dichotomous	2 yr
A319	*Jong*	*2016*	Y	Y	180	176	2.2	ITT	'Total complaints severity score' *(defined)*	Continuous	28 d
Study design 1b: NIHT + other intervention, versus other intervention alone											
A159	Taylor	2011	N	N	119	94	21.0	PP	Severity of AOM symptoms: ETG-5	Continuous	5 d
A287	Villanueva	2012	N	N	99	99	0.0	PP	Normality of weight per height	Dichotomous	12 mo
A297	Mourão	2014	N	Y	50	50	0.0	PP	Clinical attachment level	Continuous	12 mo
A304	Thinesse-Mallwitz	2015	N	Y	523	511	2.3	PP / ITT	Those with 'treatment response' *(defined)*	Dichotomous	4 d

Abbreviations: AOM, acute otitis media; ETG-5, 5-item, ear treatment group symptom questionnaire; ITT, intention to treat; PP, per protocol; RQLQ, rhino-conjunctivitis quality of life questionnaire; VAS, visual analogue scale; N, no; Y, yes.
[a]Non-inferiority trial.
[b]Equivalence trial.

not be derived, or for which only non-parametric analysis was given. For the remaining 10 studies, 6 were in study-design category 1a (other-intervention control) and 4 were in category 1b ('[A + B] versus B'); there were no trials in category 2 (no-treatment control). Three category 1a studies with extractable data were either equivalence or non-inferiority trials and were examined as a separate group.

Overall Relative Effect Sizes

Study Design 1a: Other-Intervention Control

Given merely three superiority trials of study design 1a with extractable quantitative data, and also with manifest diversity of clinical conditions (common cold and leprosy) and measured end-points (4 days to 2 years), it was deemed inappropriate to merge OR and SMD data for these studies. Individually, each of two of the trials (A150, Gassinger; A155, Maiwald) had a non-significant effect favouring homeopathy; the third trial (A305, Chakraborty) had a statistically significant effect favouring homeopathy (►**Table 4**), though the very large effect size (OR, 695) was evidently an extreme outlier.

As was shown in ►**Table 3**, each of the three trials displayed high risk of bias. Two trials (A150, Gassinger; A155, Maiwald) were equally pragmatic and explanatory in attitude; the third trial (A305, Chakraborty) was more pragmatic than explanatory.

Table 3 Risk-of-bias and external validity assessments for RCTs: seven studies excluded from meta-analysis shown in italics

			Risk-of-bias domain									
No.	First author	Year	I	II	IIIA	IIIB	IV	[a]V	VI	Risk of bias	Risk-of-bias rating	*PRECIS* assessment (total score per assessor)
Study design 1a: NIHT versus other intervention												
A150	Gassinger	1981	U	U	N	U	Y	Y	U	High	C1.4	Equally pragmatic and explanatory (30, 30)
A151	Issing	2005	U	U	Y	Y	Y	Y	Y	Uncertain	B2	Equally pragmatic and explanatory (26, 29)
A153	*Karow*	*2008*	U	U	Y	Y	Y	[b]N	U	High	C1.3	Equally pragmatic and explanatory (30, 32)
A155	Maiwald	1988	Y	Y	N	Y	N	Y	Y	High	C2.0	Equally pragmatic and explanatory (29, 29)
A156	*Shealy*	*1998*	U	U	Y	Y	Y	[b]N	Y	High	C1.2	Equally pragmatic and explanatory (32, 34)
A160	*Totonchi*	*2007*	U	U	[c]N	Y	Y	[b]N	U	High	C2.3	More pragmatic than explanatory (35, 40)
A162	Weiser	1998	Y	U	Y	Y	Y	Y	Y	Uncertain	B1	Equally pragmatic and explanatory (26, 28)
A163	Weiser	1999	U	Y	Y	Y	Y	Y	Y	Uncertain	B1	Equally pragmatic and explanatory (28, 33)
A164	*Wiesenauer*	*1987*	U	U	U	U	U	[b]N	U	High	C1.6	Equally pragmatic and explanatory (27, 27)
A285	*González de Vega*	*2013*	Y	Y	N	N	Y	[b]N	Y	High	C3.0	More explanatory than pragmatic (24, 24)
A286	*Pellow*	*2013*	U	U	U	U	Y	[b]N	Y	High	C1.4	Equally pragmatic and explanatory (27, 26)
A305	Chakraborty	2015	U	U	[c]N	N	Y	Y	U	High	C2.3	More pragmatic than explanatory (35, 41)
A319	*Jong*	*2016b*	Y	Y	N	N	Y	[b]N	Y	High	C3.0	Equally pragmatic and explanatory (29, 33)
Study design 1b: NIHT + other intervention, versus other intervention alone												
A159	Taylor	2011	Y	U	N	N	N	Y	Y	High	C3.1	Equally pragmatic and explanatory (29, 32)
A287	Villanueva	2012	Y	U	[c]N	Y	Y	Y	U	High[d]	C1.2	More pragmatic than explanatory (36, 39)
A297	Mourão	2014	Y	Y	N	Y	Y	Y	Y	High	C1.0	More explanatory than pragmatic (24, 24)
A304	Thinesse-Mallwitz	2015	Y	Y	N	N	Y	Y	Y	High	C2.0	Equally pragmatic and explanatory (29, 32)

Abbreviations: NIHT, non-individualised homeopathic treatment; RCTs, randomised controlled trials; N, no (not free of bias); U, unclear; Y, yes (free of bias).

[a]Unless a published study protocol was available, completeness of reporting was judged on correspondence of *Results* section with details in *Methods* section of the article.

[b]Data not extractable for meta-analysis.

[c]Expected high risk of bias in domain IIIA—trial assessed as 'More pragmatic than explanatory'.

[d]High risk of bias in domain IIIA only: trial is otherwise 'Uncertain risk of bias' overall.

This small number of low-quality trials did not justify our intended sensitivity analyses or our planned sub-group analyses.

Equivalence and Non-Inferiority Trials

Individually, one of the three trials (A151, Issing, which studied vertigo) had a non-significant effect favouring homeopathy; each of the other two trials (A162, Weiser [vertigo]; A163, Weiser [seasonal allergic rhinitis]) had a non-significant effect favouring control (►**Table 4**).

The pooled effect estimate for these three trials was SMD = 0.08 (95% CI: –0.13 to 0.28; $p = 0.46$) (see ►**Fig. 2**). Collectively, the trials are clinically heterogeneous (two studies on vertigo and one on seasonal allergic rhinitis) but were statistically non-heterogeneous. Each of the trials displayed uncertain risk of bias and a study attitude that was equally pragmatic and explanatory (►**Table 3**).

Due to the close similarity of trial attributes and the small number of the studies, it was deemed inappropriate to conduct sensitivity analysis or sub-group analysis.

Table 4 Effect size statistics for trials with extractable quantitative data

No.	First author	Year	'Main' outcome identified	Homeopathy	Control	Effect size (95%CI)	Direction of effect	*p*-Value
Study design Ia: NIHT versus other intervention								
A150	Gassinger	1981	Sum of symptoms	Mean = 1.9, SD = 1.23, *n* = 22	Mean = 2.31, SD = 1.34, *n* = 31	SMD = –0.31 (–0.86, 0.24)	Homeopathy	0.27
A155	Maiwald	1988	Proportion with symptom score improvement	18 of 62	12 of 53	OR = 1.40 (0.60, 3.25)	Homeopathy	0.44
A305	Chakraborty	2015	Presence of skin sensation	28 of 30	0 of 30	OR = 695 (32, 15116)	Homeopathy	< 0.001
A151	[a]Issing	2005	Frequency of vertigo episodes (per day)	Mean = 2.1, SD = 3.5, *n* = 79	Mean = 2.5, SD = 4.0, *n* = 75	SMD = – 0.11 (–0.42, 0.21)	Homeopathy	0.51
A162	[b]Weiser	1998	Frequency of vertigo episodes (per day)	Mean = 1.00, SD = 2.30, *n* = 53	Mean = 0.70, SD = 1.08, *n* = 52	SMD = 0.17 (–0.22, 0.55]	Control	0.40
A163	[b]Weiser	1999	RQLQ	Mean = 1.57, SD = 1.09, *n* = 68	Mean =1.33, SD = 1.14, *n* = 67	SMD = 0.21 (–0.12, 0.55)	Control	0.22
Study design Ib: NIHT + other intervention, versus other intervention alone								
No.	First author	**Year**	**'Main' outcome identified**	**Homeopathy**	**Control**	**Effect size (95%CI)**	**Direction of effect**	***p*-Value**
A159	Taylor	2011	Severity of symptoms: ETG-5	Mean = 4.1, SD = 8.0, *n* = 44	Mean = 2.9, SD = 7.1, *n* = 50	SMD = 0.16 (–0.25, 0.56)	Control	0.45
A287	Villanueva	2012	Normality of weight per height	42 of 50	15 of 49	OR = 11.90 (4.51, 31.39)	Homeopathy	<0.001
A297	Mourão	2014	Clinical attachment level	Mean = 4.17, SD = 0.41, *n* = 25	Mean = 4.57, SD = 0.35, *n* = 25	SMD = –1.03 (–1.63, 0.44)	Homeopathy	<0.001
A304	Thinesse-Mallwitz	2015	Those with 'treatment response' *(defined)*	40 of 259	17 of 252	OR = 2.52 (1.39, 4.58)	Homeopathy	0.002

Abbreviations: CI, confidence interval; ETG-5, 5-item, ear treatment group symptom questionnaire; OR, odds ratio; RQLQ, rhino-conjunctivitis quality of life questionnaire; SD, standard deviation; SMD, standardised mean difference.
[a]Non-inferiority trial.
[b]Equivalence trial.

Study Design 1b: '[A + B] versus B'

There were four trials in this category: two with dichotomous data, and two with continuous data, all of which were extractable for analysis. With merely four trials in total, it was inappropriate to merge OR and SMD data for these studies.

Individually, three of the four trials (A287, Villanueva; A297, Mourão; A304, Thinesse-Mallwitz) had a significant effect favouring homeopathy; the other trial (A159, Taylor) had a non-significant effect favouring control (►**Table 4**).

For the two trials with dichotomous data (A287, Villanueva; A304, Thinesse-Mallwitz), the pooled effect estimate

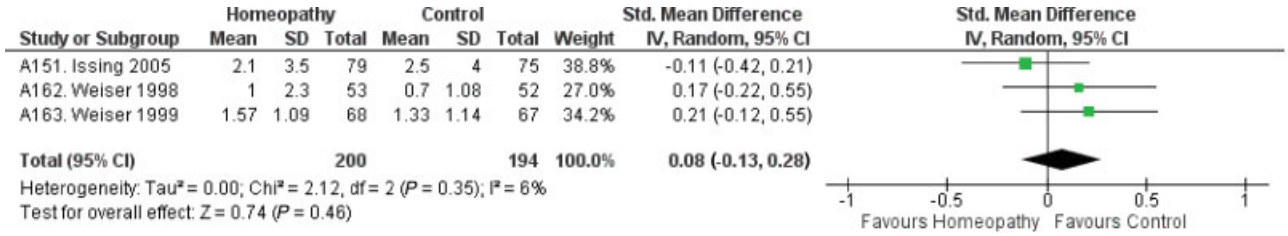

Fig. 2 Forest plot for three equivalence or non-inferiority trials: study design 1a (NIHT versus other intervention). CI, confidence interval; IV, inverse-variance method; NIHT, non-individualised homeopathic treatment; SD, standard deviation; Std, standardised.

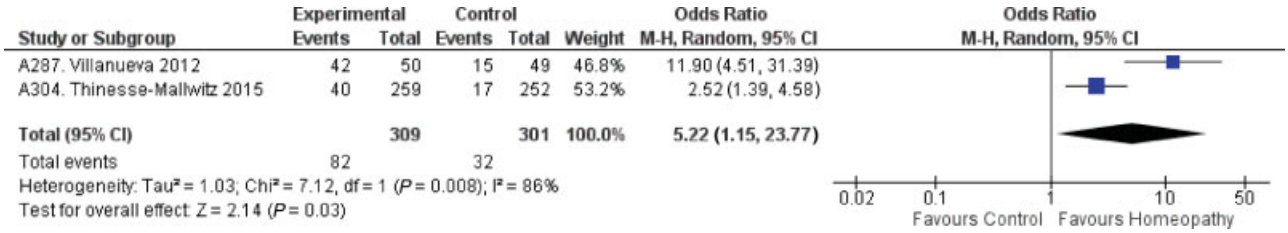

Fig. 3 Forest plot for two RCTs with dichotomous data: study design 1b ('[A + B] versus B'). CI, confidence interval; M-H, Mantel-Haenszel method; RCTs, randomised controlled trials.

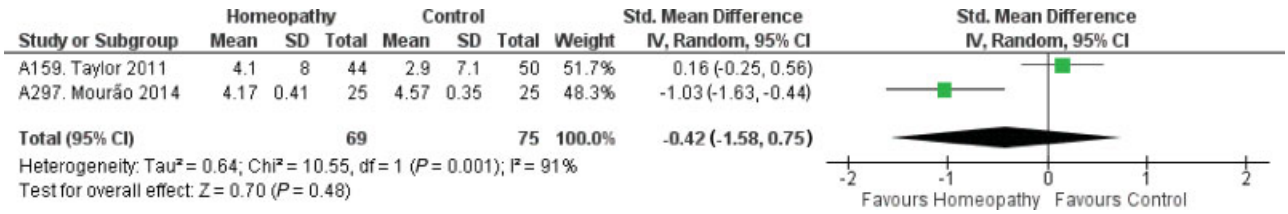

Fig. 4 Forest plot for two RCTs with continuous data: study design 1b ('[A + B] versus B'). CI, confidence interval; IV, inverse-variance method; RCTs, randomised controlled trials; SD, standard deviation; Std., standardised.

was OR = 5.22 (95% CI: 1.15 to 23.77; $p = 0.03$); $I^2 = 86\%$, indicating extreme heterogeneity (►**Fig. 3**). For A287, Villaneuva, the large effect size (OR, 11.90) was evidently an outlier. With a focus on malnutrition and upper respiratory tract infection respectively, the two studies displayed marked *clinical* heterogeneity.

For the two trials with continuous data (A159, Taylor; A297, Mourão), the pooled effect estimate was SMD = –0.42 (95% CI: –1.58 to 0.75; $p = 0.48$); $I^2 = 91\%$, indicating extreme heterogeneity (►**Fig. 4**). These two studies also displayed marked *clinical* heterogeneity (acute otitis media and chronic periodontitis respectively).

As was shown in ►**Table 3**, each of the four 1b-design trials displayed high risk of bias. The study attitudes of these trials were a mix of pragmatic and explanatory. One of the RCTs (A287, Villanueva), which was more pragmatic than explanatory, had high risk of bias in domain IIIa only: a sensitivity analysis (which we have not undertaken) would have taken that factor into account. For the same reason as for the 1a trials above, we did not undertake sensitivity or sub-group analyses.

Disease- and Category-Specific Relative Effect Sizes

Despite the overall diversity of clinical conditions and categories of condition, there were two trials that focused on *common cold* and shared the same study design (A150, Gassinger; A155, Maiwald). The study medicine in each case comprised *Eupatorium perfoliatum*; the comparator in each case was acetylsalicylic acid. Merging the OR and the SMD data for these studies, OR = 1.54 (95% CI: 0.81 to 2.93; $p = 0.19$) (►**Fig. 5**). Each study has high risk of bias; each is equally pragmatic and explanatory.

Discussion

None of the 17 studies was judged to comprise reliable evidence (i.e. there were no A- or B1*-rated trials), 14 being assessed as *high risk of bias*. Seven of the 17 articles failed to yield data suitable for meta-analysis, leaving a total of 10 studies in different categories of study design. Because of their diverse clinical nature and outcome measures—as well as their small number—it was deemed inappropriate to conduct meta-analysis on the three superiority trials of study design 1a. For the equivalence and non-inferiority trials collectively, the small, non-significant, effect size observed in meta-analysis (SMD = 0.08; $p = 0.46$) was consistent with a conclusion that NIHT did not differ from treatment by a comparator (*Ginkgo biloba* or betahistine) for vertigo or (cromolyn sodium) for seasonal allergic rhinitis. For four studies of design 1b, any significant effect favouring adjunctive NIHT was mitigated by their extreme heterogeneity. The overall low intrinsic quality and number of studies add further caution in trying to reach any

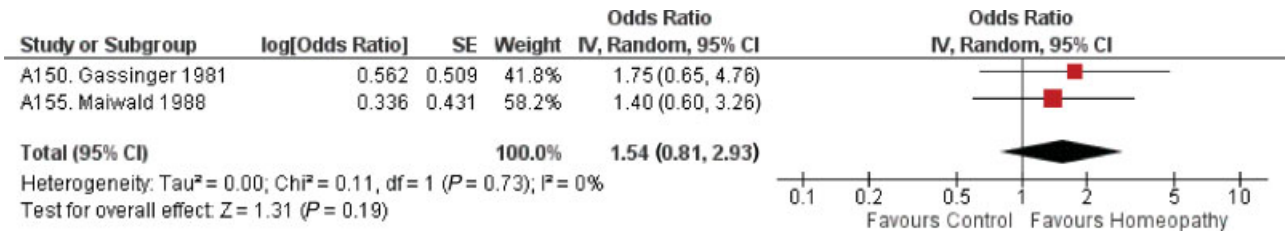

Fig. 5 Forest plot for two RCTs of NIHT: clinical condition, common cold. CI, confidence interval; IV, inverse-variance method; NIHT, non-individualised homeopathic treatment; RCTs, randomised controlled trials; SD, standard deviation.

meaningful conclusions about comparative effectiveness of NIHT. The most recent NIHT literature (from 2017) provides just one further OTP-controlled trial of NIHT[24] that appears to satisfy our SR inclusion criteria.

Lack of replication of subject matter allowed just one meta-analysis by clinical condition: pooling data from two 'C'-rated trials of common cold, where NIHT comprised *Eupatorium perfoliatum* in each case, found no difference in treatment effect compared to that of aspirin (OR = 1.54; $p = 0.19$). It is thus inconclusive whether *Eupatorium* and acetylsalicylic acid (aspirin) differ in effectiveness for treating common cold symptoms.

Within the pragmatic—explanatory continuum, the internal validity for pragmatic study design would naturally be lower than that for explanatory design.[8] Logically, therefore, an intrinsically low-quality study might be expected to have more emphasis placed on external validity; however, it is notable that, despite their OTP-controlled design, only three of the studies possessed a truly pragmatic study attitude. The external validity of 14 trials was, therefore, little or no higher than that of an overtly explanatory trial, with limited generalisability of findings to a wider population of patients. Being equally pragmatic and explanatory in attitude, the equivalence and non-inferiority trials of NIHT (A151, Issing; A162, Weiser; A163, Weiser) were found to lack clear external validity; moreover, their overall risk of bias in each case was uncertain ('B1'- or 'B2'-rated) rather than high. Amongst the studies that examined adjunctive homeopathy, just one (A287, Villanueva) of the four was more pragmatic than explanatory in attitude. The latter was the solitary example of a trial with high risk of bias in domain IIIa only, preventing our intended sensitivity analysis on this point. We have previously commented on the rigour of the Cochrane approach in this respect.[2]

Although the *PRECIS* tool was originally developed to help at the *design* stage of clinical trials,[10] we found it fairly straightforward to apply in assessing the pragmatic/explanatory attitude of a completed study. The scoring notation we implemented proved able to approximate a given trial's positioning on the pragmatic—explanatory continuum, with little difficulty in reconciling inter-rater assessments. The 10-domain judgmental criteria were operationalised successfully, though there was often a lack of information in the published articles to form a clear opinion on the following: practitioner expertise in homeopathy (domain 3); practitioner expertise in the comparison intervention (domain 5); participant compliance (domain 8); protocol adherence (domain 9).

Conclusion

It is currently not possible to form a decisive conclusion regarding comparative effectiveness of NIHT on health-related outcomes, assessed within the context of OTP-controlled trials. Generalisability of findings is limited by the very small number of studies that have predominantly pragmatic attitude. Amongst NIHT studies currently, the highest intrinsic quality is seen in those designed as an equivalence or non-inferiority trial. Mirroring the conclusion of our recent SR,[2] future OTP-controlled trials in NIHT should aim, as far as possible, to promote both internal validity and external validity.

Highlights

- This systematic review focuses on randomised controlled trials (RCTs) of non-individualised homeopathic treatment (NIHT) in which the control (comparator) group was other than placebo.
- For each eligible trial, risk of bias was assessed using Cochrane methods, and its relative pragmatic or explanatory attitude was approximated using the *PRECIS* tool.
- Seventeen RCTs, representing 15 different medical conditions, were eligible for inclusion.
- Fourteen RCTs were rated 'high risk of bias'; the other three trials were rated 'uncertain risk of bias'.
- Only three RCTs were judged to have clearly pragmatic study attitude.
- Quantitative data extraction did not yield a decisive conclusion about the comparative effectiveness of NIHT.

Supplementary Files

Supplementary File 1: PRISMA 2009 checklist.

Supplementary File 2: Details of records: OTP-controlled RCTs of NIHT.

Supplementary File 3: Risk-of-bias summary graph.

Authors' Contributions

RTM devised and led the study, developed the study protocol and contributed to all facets of the work. YYYF helped to develop the study protocol, co-assessed trials for risk of bias, co-assessed trials for pragmatic/explanatory study attitude, contributed to data interpretation and edited the manuscript. PV helped to develop the study protocol, co-assessed trials for pragmatic/explanatory study attitude, contributed to data interpretation and edited the manuscript. AKLT co-assessed trials for risk of bias and contributed to data interpretation. JRTD helped to develop the study protocol, contributed to data interpretation and edited the manuscript. All authors have approved the final manuscript.

Competing Interests

Authors RTM, YYYF, PV and AKLT are (or were) associated with a homeopathy organisation whose significant aim is to clarify and extend an evidence base in homeopathy. RTM holds an independent research consultancy contract with the Deutsche Homöopathie-Union, Karlsruhe, Germany. YYYF and AKLT belong to Living Homeopathy Ltd., which has contributed funding to some (but not this current) HRI project work. RTM and PV have no other relationships or activities that could appear to have influenced the submitted work. JRTD had no support from any organisation for the submitted work; in the last 3 years, and for activities outside the submitted study, he received personal fees, royalties or out-of-pocket expenses for advisory work,

invitational lectures, use of rating scales, published book chapters or committee membership; he receives royalties from Springer Publishing Company for his book, *A Century of Homeopaths: Their Influence on Medicine and Health.* JTRD has no other relationships or activities that could appear to have influenced the submitted study.

Funding
The systematic review programme has been supported by a grant from the Manchester Homeopathic Clinic.

Acknowledgements
We thank Daniela Hacke, Karl und Veronica Carstens-Stiftung, Essen, Germany, for assistance with the literature search and retrieval of relevant full-text articles.

References

1 Swayne J. International Dictionary of Homeopathy. Edinburgh: Churchill Livingstone; 2000
2 Mathie RT, Ulbrich-Zürni S, Viksveen P, et al. Systematic review and meta-analysis of randomised, other-than-placebo controlled, trials of individualised homeopathic treatment. Homeopathy 2018;107:229–243
3 Mathie RT, Lloyd SM, Legg LA, et al. Randomised placebo-controlled trials of individualised homeopathic treatment: systematic review and meta-analysis. Syst Rev 2014;3:142
4 Mathie RT, Ramparsad N, Legg LA, et al. Randomised, double-blind, placebo-controlled trials of non-individualised homeopathic treatment: systematic review and meta-analysis. Syst Rev 2017;6:63
5 Ernst E. Classical homeopathy versus conventional treatments: a systematic review. Perfusion (Nürnberg) 1999;12:13–15
6 Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. Trials 2009; 10:37
7 Sedgwick P. External and internal validity in clinical trials. BMJ 2012;344:e1004
8 Sedgwick P. Explanatory trials versus pragmatic trials. BMJ 2014; 349:g6694
9 Bornhöft G, Maxion-Bergemann S, Wolf U, et al. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. BMC Med Res Methodol 2006;6:56
10 Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol 2009;62:464–475
11 Mathie RT, Fok Y, Viksveen P, et al. Systematic review and meta-analysis of randomised, other-than-placebo (OTP) controlled, trials of individualised homeopathic treatment (IHT): Study protocol. Version 1.2; 25 January 2018. Available at: https://www.hri-research.org/hri-research/learning-more-from-existing-evidence/systematic-review-programme/
12 Mathie RT, Hacke D, Clausen J, Nicolai T, Riley DS, Fisher P. Randomised controlled trials of homeopathy in humans: characterising the research journal literature for systematic review. Homeopathy 2013;102:3–24
13 Baker DG, Myers SP, Howden I, Brooks L. The effects of homeopathic *Argentum nitricum* on test anxiety. Complement Ther Med 2003;11:65–71
14 Higgins JPT, Altman DG. Assessing risk of bias in included studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions; Version 5.1.0. The Cochrane Collaboration; 2011. Chapter 8
15 Linde K, Clausius N, Ramirez G, et al. Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials. Lancet 1997;350:834–843
16 Shang A, Huwiler-Müntener K, Nartey L, et al. Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. Lancet 2005;366:726–732
17 Hewitt CE, Torgerson DJ, Miles JNV. Is there another way to take account of non-compliance in randomized controlled trials? Can Med Assoc J 2007;175:347–348
18 Higgins JPT, Deeks JJ, Altman DG. Special topics in statistics. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions; Version 5.1.0. The Cochrane Collaboration; 2011. Chapter 16
19 Sedgwick P. Equivalence trials. BMJ 2013;346:f184
20 Sedgwick P. What is a non-inferiority trial? BMJ 2013;347: f6853
21 Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions; Version 5.1.0. The Cochrane Collaboration; 2011. Chapter 9
22 Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med 2000;19:3127–3131
23 Schünemann HJ, Oxman AD, Vist GE, et al. Interpreting results and drawing conclusions. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions; Version 5.1.0. The Cochrane Collaboration; 2011. Chapter 12
24 Palm J, Kishchuk VV, Ulied À, et al; TocTo Research Group. Effectiveness of an add-on treatment with the homeopathic medication SilAtro-5-90 in recurrent tonsillitis: An international, pragmatic, randomized, controlled clinical trial. Complement Ther Clin Pract 2017;28:181–191