Systematic review and meta-analysis of randomised, other-than-placebo (OTP) controlled, trials of *non*-individualised homeopathic treatment (NIHT):

Study protocol

[1] **Robert T Mathie**

[2] **Yvonne Fok**

[3] **Petter Viksveen**

[1] **E Rachel Roberts**

[2] **Aaron To**

[4] **Lynn A Legg**

[5] **Jonathan R T Davidson**

[1] Homeopathy Research Institute (HRI), London, UK

[2] The Hong Kong Association of Homeopathy

[3] Department of Health Studies, University of Stavanger, Norway

[4] Department of Biomedical Engineering, University of Strathclyde, Glasgow, UK

[5] Department of Psychiatry and Behavioral Science, Duke University Medical Center, Durham, North Carolina, USA

Version 1.2

25 January 2018

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. The prescription of a homeopathic medicine ensues from an individual's consultation with a practitioner. It is believed that the effect of administering a homeopathic medicine is to stimulate a healing response in the patient.[1]

There are several distinct forms of homeopathy, the main types being 'individualised' (or 'classical') homeopathy, 'clinical' homeopathy, 'complex' homeopathy, and isopathy:

In **individualised homeopathy** – as originally defined by its founder, Samuel Hahnemann – typically a single homeopathic medicine is selected on the basis of the 'total symptom picture' of a patient, including his/her mental, general and constitutional type.

In *clinical homeopathy*, 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 these three approaches involves matching a patient with the 'total symptom picture' of an individually prescribed homeopathic medicine: each is thus termed **non-individualised homeopathy**.

To date, our main systematic review focus has been individualised or non-individualised homeopathy in the context of *placebo-controlled* RCTs.[2,3,4] The current protocol focuses solely on RCTs of **non-individualised homeopathic treatment (NIHT)**, in which the control (comparator) group is something ***other than placebo (OTP)*** – they can be regarded as 'comparative effectiveness' studies.

Two essentially different comparator 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.

The nature of the research evidence in homeopathy has long been a matter of scientific debate. Recently, however, the argument has reached the point of impasse, despite pleas to the contrary.[5] Further clarification of the existing evidence, based on explicit, transparent and replicable systematic review methods, might moderate such divergent opinions: the nature and the quality of the research evidence must be considered, with an additional requirement to reflect the distinction between individualised and non-individualised homeopathy.

The pinnacle of the hierarchy of clinical research publications ('type 1' evidence) comprises systematic reviews (SRs), of which several have been published on RCTs in homeopathy. Some SRs have focused on specific medical conditions, typically examining placebo-controlled trials, with conclusions about efficacy of homeopathic medicines that are variously positive,[e.g. 6,7,8] negative[e.g. 9,10,11] or non-conclusive.[e.g. 12,13,14]

Five 'global', or 'comprehensive', SRs have examined the RCT research literature on homeopathy in general, including the broad spectrum of medical conditions that have been researched, and by all forms of homeopathy. Four of these SRs reached the cautious conclusion that, overall, the effect of a homeopathic medicine differs from that of placebo.[15,16,17,18] When Linde and colleagues carried out a sensitivity analysis on the data that informed their 1997 global SR based on quality of the 89 eligible trials, the observed effects were substantially reduced, though homeopathic medicines remained significantly more effective than placebo for all but the final five highest-quality trials collectively.[19] Neither of Linde's reviews found sufficient evidence to draw clear conclusions about the efficacy of homeopathic medicines for any specific medical condition. The SR by Shang et al, published in 2005, concluded that there was "weak evidence for a specific effect of

homoeopathic remedies…compatible with the notion that the clinical effects of homoeopathy are placebo effects".[20] Shang's methods and conclusions have subsequently been criticised.[e.g. 21]

No previous global 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:[22] published 18 years ago, it identified six eligible trials, two of which favoured homeopathy, two favoured conventional drugs, and two were non-conclusive. It was concluded overall that the 'value of individualized homoeopathy relative to allopathic treatments' was not known.

Previous reviews contain two key limitations:

1. Global SRs have typically assessed the RCT findings of all forms of homeopathy (individualised, clinical, complex, isopathy) together, as if they are the same intervention. As discussed above, there are marked differences in the nature of the therapeutic interventions, and the distinction between them is important, for it affects the interpretation of the research findings in each case. In a placebo-controlled trial of individualised homeopathy, 'efficacy' applies to the range of homeopathic medicines prescribed to the individuals included in the trial.[4] A placebo-controlled RCT of a particular homeopathic medicine (non-individualised homeopathy), however, allows conclusions about that medicine's efficacy for the clinical condition investigated in the cohort of subjects concerned.

2. Though not *systematic* reviews, some previous accounts of homeopathy research, including our own,[23] have summarised the findings of RCTs using 'vote counting', whereby each trial is designated 'positive' or 'negative' or 'non-conclusive' based on its most important statistical findings. While such an approach has the advantage that it overcomes problems associated with heterogeneous groups of trials and reflects the condition-specific nature of the research evidence, it does not grapple adequately with the key matters of treatment effect size or direction, or the strength of the evidence. For an OTP-controlled trial, vote counting might lead to 'no significant difference between test intervention and active comparator' being interpreted as positive evidence for the test intervention, when in fact statistical equivalence can be concluded only for a trial whose hypothesis is based on equivalence rather than superiority.[24] Nor does this 'vote counting' method reflect a single 'main outcome measure' of each trial in a systematic way. There is a need to quantify treatment effects of homeopathic interventions, and the use of a systematically and consistently determined 'main outcome measure' per RCT is helpful in focusing on matters of greatest clinical importance for a given medical condition (or category of conditions).

Four additional matters also need to be addressed:

a. Nearly all SRs to date have examined RCTs of treatment and of prophylaxis indistinguishably. However, the homeopathic approach to these two clinical scenarios is different: a person's *actual* symptoms are the target of homeopathic *treatment*, but the *expected* or *typical* symptoms are the basis for homeopathic *prophylaxis*.

b. The *internal validity* of a trial (the extent to which the design, conduct and analysis has minimised or avoided biases in its comparison of treatments[25,26]) reflects the quality of its methods of randomisation, blinding, and a number of other key attributes. Some comprehensive reviews have used a numerical system, such as the Jadad score,[27] to assess RCT quality in homeopathy. More modern systems of assessment do not allocate single overall scores; instead, they adopt qualitative standards against which a trial's internal validity is judged as having low, uncertain or high risk of bias.[28] Neither system is intended to enable the identification of finer distinctions in degree of quality.

c. Concerns about RCT quality in homeopathy go beyond – though include – internal validity (risk of bias).[29] Indeed, in the context of OTP-controlled RCTs, it is the *external validity* (generalisability of results) that should normally be maximised,[30] though it is seldom formally addressed in SRs of such trials.[31] A tool (*PRECIS*), developed by Thorpe et al. in 2009 for use in SRs and their design, is available to assess a trial's positioning on the 'explanatory–pragmatic continuum'.[32] For our

SR of OTP-controlled trials, we shall assess each trial's internal validity and also the extent to which it is 'pragmatic' in approach; the model validity of the same trials may be addressed separately in a subsequent project.

d. Few of the previous SRs in homeopathy have made the distinction between substantive and minor research articles or between the peer-reviewed and non-peer-reviewed research literature: an abstract presented at a conference, for example, has often been given a status equal to that of a paper published in a high-ranking academic journal.[e.g. 17,20] Peer review is an important, though by no means flawless, surrogate for research quality. For some commentators, it is "an essential arbiter of scientific quality" and "information about the status of research results is as important as the findings themselves".[33] SRs in homeopathy need to reflect, *a priori*, the distinction between the substantive peer-reviewed journal literature and other, lesser, categories of research evidence.

Aim of the study

Our overarching objective is to examine comparative effectiveness of NIHT in OTP-controlled trials of any clinical condition in adults or children. Using meta-analysis, we aim to evaluate RCTs that have investigated NIHT: (1a) in comparison to another therapeutic intervention; (1b) adjunctively with another treatment intervention in comparison to the other intervention alone ('[A+B] versus B'); or (2) compared with no other intervention. A single 'main outcome measure' will be identified per RCT.

An additional aim, if feasible, is to evaluate comparative effectiveness of NIHT for any clinical condition or category of conditions for which there was >1 eligible RCT. In all cases, we shall reflect matters of internal validity (risk of bias) and external validity (pragmatic/ explanatory study attitude).

METHODS

Eligibility criteria, information sources, study selection and data collection

The eligible research literature was identified, to *PRISMA* standards, in a previous paper by our group.[34] From 489 potentially eligible records found up to and including 2011 (fully up-to-date at the time), 263 fulfilled the criteria of a substantive, non-repeat, journal paper that reported a randomised and controlled study of homeopathy.

Fifteen of those records reported an OTP-controlled trial of non-individualised homeopathy and were published in the peer-reviewed journal literature. Subsequent search updates – most recently in March 2017 – have identified a total of **11** additional relevant papers published in 2012–2016. **Figure 1** is based on our original *PRISMA* flowchart,[34] and incorporates these 11 papers, giving **26** in total. Further details of the relevant search update are included in the HRI website.[35]

Specific **exclusion criteria** have been applied, as appropriate, to these 26 records:

- Trials of homeopathic prophylaxis[a]
- Trials with crossover design[b]
- Research using radionically prepared 'homeopathic' medicines [e.g. 36]
- The tested intervention is NIHT in tandem with other (complementary or conventional) medicine or therapy, and where the nature of the comparator intervention makes it impossible to distinguish any effects solely due to NIHT.[c]
- Other specified reason.

Whereas a placebo-controlled trial of non-individualised homeopathy can be fully

[a] *Prophylaxis:* A trial on healthy individuals in which the homeopathic intervention aims to prevent the occurrence of disease *de novo* (i.e. 'primary prevention'). Studies using a strategy of primary prevention, with subsequent treatment as necessary, are categorised 'treatment' trials.
Treatment: A trial in which the first homeopathic intervention takes places after the onset of symptoms associated with disease. Studies on sub-clinical disease or the control of recurrent disease ('secondary prevention') are categorised 'treatment' trials. RCTs of homeopathic prophylaxis will be appraised in a separate SR.

[b] In due course, crossover trials will be appraised separately from those of parallel-group design.

[c] This study design is distinct from the eligible 'A versus [A + B]' design, and from eligible studies that allow concomitant conventional medication to remain ongoing in the subjects of *each* study group. See *Appendix 1* for tabulation of eligible and ineligible study designs.

blinded, it is more difficult – and sometimes impossible – to achieve such blinding in a corresponding OTP-controlled trial. Unlike the case for our systematic review of placebo-controlled trials, therefore, patient- and practitioner-unblinded trials *will* be eligible for the current SR of OTP trials. Such trials are likely to be rated 'high risk of bias' in the relevant assessment domain: 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.

Nine records met the above exclusion criteria, leaving **17** therefore that are eligible for SR – see Figure 1. Any record whose main outcome measurement is not extractable (see below) will be ineligible for meta-analysis.

Only published data will be eligible for analysis. Because it is recognised that contacting the original authors of RCTs may lead to limited or overly positive answers,[28] the authors of eligible papers will not be approached for clarification on unclear or missing facets of any of their methods or results; however, original authors' cross-reference to their previously published study methods will be followed up and taken into account as necessary. For trials with more than two study groups, and where such trials have not previously been catalogued under 'placebo-controlled', only the data concerning comparisons between NIHT and OTP will be extracted from the 17 papers; in relevant cases of more than one OTP control, a study group comprising actual treatment will be favoured for analysis over one comprising 'no treatment'.

Study characteristics and data items

Two reviewers independently will extract relevant data using a standard data recording approach, in spreadsheet format (Microsoft *Excel*). The data extracted per trial will include, as appropriate: demographics of participants (gender, age range, medical 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 will be: sample size and missing data (including patient non-compliance) for each intervention group; whether power calculation carried out; whether intention-to-treat (ITT), per-protocol, complier-average-causal-effect (CACE),[37] or other type of primary analysis.

Identification of 'main outcome measure' per RCT:

For each trial, and for the purposes of risk-of-bias assessment, we shall identify a single 'main outcome measure' using a refinement of the approaches adopted by Linde et al. and by Shang et al.[17,20] Each trial's 'main outcome measure' will be identified based on the following hierarchical ranking order (consistent with the WHO ICF Classification System for Levels of Functioning Linked to Health Condition):[d]

- Mortality
- Morbidity
    - Treatment failure
    - Pathology; symptoms of disease
- Health impairment (loss/abnormality of function, incl. presence of pain)
- Limitation of activity (disability, incl. days off work/school because of ill health)
- Restriction of participation (quality of life)
- Surrogate outcome (e.g. blood test data, bone mineral density).

We shall follow the WHO ICF system regardless of what measure may have been identified by the investigators as their 'primary outcome'. In cases where, in the judgment of the reviewers, there are two or more outcome measures of equal greatest importance within the WHO ICF rank order, the designated 'main outcome measure' will be selected randomly from those two or more options using the toss of coins or dice.

Unless otherwise indicated, the single end-point (measured from the start of the intervention) associated with the designated 'main outcome measure' will be taken as the last follow-up at which data are reported for that outcome.

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

Assessing risk of bias (internal validity)

Using the standard criteria defined by Cochrane,[28] the extraction of information will enable 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;[e] (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 assessors will mutually scrutinise and compare their judgments, with discrepancies between them resolved by consensus discussion and, if necessary, the input of a third assessor. A risk-of-bias summary table will be produced, characterising each of the 17 eligible records. For domain IV, a trial will normally be regarded as no better than 'unclear' if there is greater than 20% participant attrition rate, irrespective of whether ITT analysis has been carried out. Domain V will automatically be designated 'high risk of bias' if its main outcome measure cannot be extracted to enable calculation of 'relative effect size' (see below). Assessment of domain VI will explicitly include appraisal of inter-group data imbalance at baseline; the source of any research sponsorship will be 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 will be designated: **low risk of bias** for all key domains; **uncertain risk of bias** for one or more key domains; **high risk of bias** for one or more key domains.[28] We shall use our novel method of nomenclature, based on the Cochrane approach, for rating risk-of-bias characteristics across all domains per trial:[2,3]

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 is designated *reliable evidence*. A 'B'-rated trial is *reliable evidence* if the uncertainty in its risk of bias is for one of domains IV, V or VI only (it is free of bias for each of domains I, II, IIIA and IIIB).

Assessing pragmatic/explanatory study attitude (external validity):

Equating external validity to study attitude, we shall adopt the *PRECIS* approach[32] to assess each trial's positioning on the pragmatic–explanatory continuum, taking account of ten domains:

1. Participant eligibility criteria;
2. Experimental intervention flexibility;
3. Experimental intervention practitioner expertise;
4. Comparison intervention;
5. Comparison intervention practitioner expertise;
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,[32] we shall aim to assess each of the ten attributes as 'more explanatory than pragmatic' or 'more pragmatic than explanatory'.

Summary measures for 'main outcome'

A 'summary of findings' table (containing relevant [per-protocol] data from the trials) and a summary risk-of-bias table will be prepared.

For the 17 relevant records of NIHT, we shall examine: (1) **overall relative effect size**; (2) **disease-specific relative effect sizes**; (3) **disease category-specific relative effect sizes**. In each of these three cases, 'relative effect size' will be taken as the difference (if relevant

[e] Domains are designated $IIIa_1$, $IIIa_2$ and IIIb to reflect their common – but separately identifiable – focus on matters connected with blinding.

– see below) between the homeopathy and the OTP groups at our pre-determined end-point of the trial, and using per-protocol data:

- For **dichotomous measures**: odds ratio (OR), with 95% confidence interval (CI);[f]
- For **continuous measures**: standardised mean difference (SMD), with 95% CI.[g]

In trials where the main outcome measure is a continuous variable, and where there are insufficient data presented to identify the mean and/or the SD per group at the defined end-point, the necessary data will be calculated or estimated, if possible, by imputing relevant other data (e.g. SD at baseline) from the same study.[38]

If the original paper does not provide or inform adequate data on the selected 'main outcome measure' to enable extraction or calculation of mean and/or SD, we shall describe the selected main outcome as 'not estimable': an alternative, estimable, outcome will *not* be sought.

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

- Those that present non-parametric data only, and where there is no information that enables the data distribution to be assessed;
- Those from which the necessary data cannot be extracted (not provided or uninterpretable).

Statistical interpretation

Interpretation of statistical finding: P≤0.05 (direction of effect toward homeopathy):

> *(1a) Other-intervention control:* NIHT is more effective than the other intervention (*but see comments below*);
> *(1b) '[A+B] versus B':* NIHT + other intervention is more effective than the other intervention alone;
> *(2) No-treatment control:* NIHT is more effective than no intervention.

Interpretation of statistical finding: P≤0.05 (direction of effect toward control):

> *(1a) Other-intervention control:* NIHT is less effective than the other intervention;
> *(1b) '[A+B] versus B':* NIHT + other intervention is less effective than the other intervention alone;
> *(2) No-treatment control:* NIHT is ineffective.

Interpretation of statistical finding: P>0.05 (direction of effect toward either homeopathy or control):

> *(1a) Other-intervention control:* Inconclusive whether NIHT and the other intervention differ in effectiveness;
> *(1b) '[A+B] versus B':* Inconclusive whether the effectiveness of NIHT + other intervention differs from that of the other intervention alone;
> *(2) No-treatment control:* NIHT is probably ineffective.

Detailed consideration of the above will reflect, where feasible, whether the other intervention is recognised as best standard care for the relevant medical condition. In cases where the comparator is best effective standard care, NIHT (either alone or as adjunctive treatment) is not expected to be found *more* effective statistically, and so our conclusions will instead reflect on matters of equivalence or non-inferiority.[24,39] This judgmental approach will predominate, by default, in the examination of trials that are prospectively attributed 'equivalence' or 'non-inferiority' by their original authors. In the latter cases, we shall reflect the original authors' pre-stated margin of equivalence or non-inferiority, as appropriate. In cases where the comparator is *not* a standard treatment for the medical condition, a clear judgment about NIHT's comparative effectiveness may not be possible.

For any RCT or group of RCTs on a given medical condition/category, the interpretation of NIHT as 'effective', 'ineffective' or 'inconclusive' will apply solely to the particular medical condition/category being examined. We recognise also that we shall be using per-protocol data (as opposed potentially to the original authors' use of ITT or CACE, for example), and that the selected main

[f] OR > 1 favours homeopathy.

[g] SMD < 0 favours homeopathy.

outcome measure/s might be insensitive to change.

Synthesis of quantitative results (if the extracted data allow)

1) Overall 'relative effect size' of NIHT

For groups of eligible RCTs that have compared NIHT (1a) with another intervention, or (1b) adjunctly with another intervention, or (2) with no treatment, the 'main outcome' data will be synthesised for meta-analysis in two separate sets of studies as appropriate: (1) using the OR of each relevant trial; (2) using the SMD of each relevant trial.[40] A summary measure of 'relative effect size' will be identified across all included studies for each of those two sets. Based on the assumption of clinical heterogeneity, the 'random effects' statistical model will be used rather than the 'fixed effect' model.[39] Illustration of findings will be by means of forest plots.

For each of study designs (1a), (1b) and (2), data – if sufficient in number – from the two sets of studies above (OR and SMD) will then also be combined into a single forest plot, re-expressing SMDs by transformation to OR, using an approximation method proposed by Chinn[41] and recommended by the Cochrane Statistical Methods Group.[39]

2) Disease-specific 'relative effect size' of NIHT

For each specific medical condition or category of conditions, for each of study designs (1a), (1b) and (2), and for which there is >1 RCT of given type and with extractable main outcome, the data will be synthesised using meta-analysis methods. For each of these particular analyses, *a single* 'main outcome measure' will be designated, if possible, for each medical condition, and reflecting the WHO classification ranking approach (see above). A summary estimate of 'relative effect size' per condition, with 95% CI and *P* value, will be illustrated by means of forest plot. The 'random effects' statistical model will again be used.[39]

3) Heterogeneity and asymmetry:

The I^2 statistic will be used to assess the variability between studies: it can take values between 0% (all of the variability is due to sampling error) and 100% (all variability is due to true heterogeneity between studies). Funnel plots will be used to assess the impact of publication bias.

Additional quantitative analyses on overall 'relative effect size' of NIHT *(specified prior to data analysis)*

Sensitivity analyses:

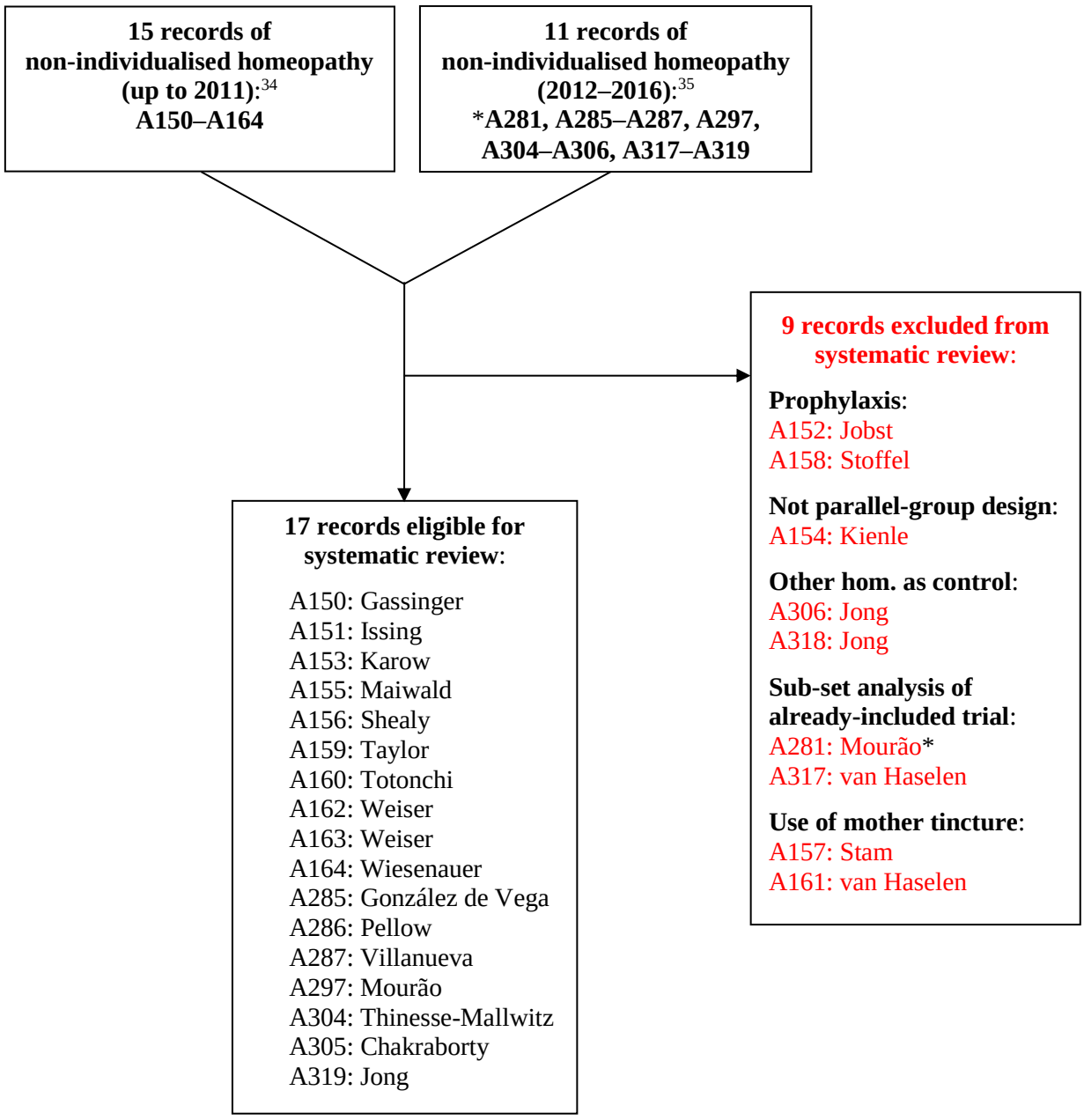
We shall carry out sensitivity analyses based separately on (1) our risk-of-bias ratings and on (2) our assessments of external validity. For a trial categorised as 'more pragmatic than explanatory', we shall accommodate an expectation of high risk of bias in domain IIIa (concerning blinding of participants and/or study personnel), which we recognise as a standard feature of pragmatic trial design in NIHT.

Sub-group analyses:

Comparative forest plots are 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 $\geq$ 12C;
- Whether or not the research sponsor is an organisation (e.g. homeopathic pharmacy) that potentially has vested interest in the trial findings.
- Whether 'clinical' homeopathy, 'complex' homeopathy, or isopathy.

Figure 1: *Details of numbered references for OTP trials, as per original* PRISMA *flowchart (2011),*[34] *and updated references published in 2012–2016.*



* A281: Mourão (2013) was originally designated 'individualised homeopathy' (see *Additional file 2* of Mathie et al. 2014).[4] Its more detailed scrutiny re-designated it 'non-individualised homeopathy': it is thus included in the present study.

References for Figure 1:
(Red font: Excluded from current systematic review)

A150 Gassinger CA, Wünstel G, Netter P (1981). Klinische Prüfung zum Nachweis der therapeutischen Wirksamkeit des homöopathischen Arzneimittels Eupatorium perfoliatum D2 (Wasserhanf composite) bei der Diagnose Grippaler Infekt. [A controlled clinical trial for testing the efficacy of the homeopathic drug Eupatorium perfoliatum D2 in the treatment of common cold]. *Arzneimittel Forschung*; **31**: 732–736.

A151 Issing W, Klein P, Weiser M (2005). The homeopathic preparation Vertigoheel versus Ginkgo biloba in the treatment of vertigo in an elderly population: a double-blinded, randomized, controlled clinical trial. *Journal of Alternative and Complementary Medicine*, **11**: 155–160.

A152 Jobst D, Altiner A, Wegscheider K, Abholz H-H (2005). Helfen intramusculäre Eigenblutgaben bei chronisch rezidivierenden Infekten der Atemwege? – Fußangeln auf dem Weg einer randomisierten Studie [Do autologous blood injections help against relapsing upper respiratory infections? - Traps on the way of doing a randomised trial]. *Zeitschrift für Allgemeinmedizin*; **81**: 258-263.

A153 Karow J-H, Abt H-P, Fröhling M, Ackermann H (2008). Efficacy of Arnica montana D4 for healing of wounds after Hallux valgus surgery compared to diclofenac. *Journal of Alternative and Complementary Medicine;* **14**: 17–25.

A154 Kienle G (1973). Wirkung von Carbo Betulae D6 bei respiratorischer Partialinsuffizienz [Effect of Carbo betulae D6 in partial respiratory insufficiency]. *Arzneimittel-Forschung/Drug Research*; **23**: 840-842.

A155 Maiwald VL, Weinfurtner T, Mau J, Connert WD (1988). Therapie des grippalen Infekts mit einem homöopathischen Kombinationspräparat im Vergleich zu Acetylsalycilsäure. Kontrollierte, randomisierte Einfachblindstudie. [Treatment of common cold with a combination homeopathic preparation compared with acetylsalicylic acid. A controlled, randomized single-blind study]. *Arzneimittel Forschung*; **38**: 578–582.

A156 Shealy CN, Thomlinson RP, Cox RH, Borgmeyer RN (1998). Osteoarthritic pain: a comparison of homeopathy and acetaminophen. *American Journal of Pain Management*; **8**: 89–91.

A157 Stam C, Bonnet MS, van Haselen RA (2001). The efficacy and safety of a homeopathic gel in the treatment of acute low back pain: a multi-centre, randomized, double-blind comparative clinical trial. *British Homeopathic Journal*, **90**: 21–28.

A158 Stoffel L, Zimmermann D, Hunkeler R, Zimmermann C, Ramos M, Fathi-Torriani M, Kropf G, Nelle M (2007). Euphrasia-Augentropfen bei Neugeborenen: Ein Pilotprojekt. [Euphrasia eye drops in neonates: a pilot project]. *Schweizerische Zeitschrift für Ganzheitsmedizin;* **19**: 254–259.

A159 Taylor JA, Jacobs J (2011). Homeopathic ear drops as an adjunct to standard therapy in children with acute otitis media. *Homeopathy*; **100**: 109-115.

A160 Totonchi A, Guyuron B (2007). A randomized, controlled comparison between arnica and steroids in the management of postrhinoplasty ecchymosis and edema. *Plastic and Reconstructive Surgery*; **120**: 271-274.

A161 van Haselen RA, Fisher PAG (2000). A randomized controlled trial comparing topical piroxicam gel with a homeopathic gel in osteoarthritis of the knee. *Rheumatology*; **39**: 714–719.

A162 Weiser M, Strösser W, Klein P (1998). Homeopathic vs. conventional treatment of vertigo: a randomized double-blind controlled clinical study. *Archives of Otolaryngology – Head and Neck Surgery*, **124**: 879–885.

A163 Weiser M, Gegenheimer LH, Klein P (1999). A randomized equivalence trial comparing the efficacy and safety of Luffa comp.-Heel nasal spray with cromolyn sodium spray in the treatment of seasonal allergic rhinitis. *Forschende Komplementärmedizin und Klassische Naturheilkunde*, **6**: 142–148.

A164 Wiesenauer M, Gaus W (1987). Orthostatische Dysregulation. Kontrollierter Wirkungsvergleich zwischen Etilefrin 5 mg und dem homöopathischen Arzneimittel Haplopappus D2. [Orthostatic dysregulation. Controlled equivalence trial with Etilefrin 5 mg vs. the homeopathic medicine Haplopappus D2]. *Zeitschrift für Allgemeinmedizin*; **63**: 18–23.

A281 Mourão LC, Moutinho H, Canabarro A (2013). Additional benefits of homeopathy in the treatment of chronic periodontitis: A randomized clinical trial. *Complement Ther Clin Pract*; **19**: 246–250.

A285 González de Vega C, Speed C, Wolfarth B, González J (2013). Traumeel vs. diclofenac for reducing pain and improving ankle mobility after acute ankle sprain: A multicentre, randomised, blinded, controlled and non-inferiority trial. *Int J Clin Pract*; **67** :979–989.

A286 Pellow J, Swanepoel M (2013). A randomised pilot study on the efficacy of milking cream and a homeopathic complex topical cream on diaper dermatitis. *Health SA Gesondheid*; **18**; #680.

A287 Villanueva DFD, Rodríguez AP, García LRG, Osés CAM (2012). Use of homeopathic formula in malnourished children. Int J High Dilution Res; **11**: 25–32.

A297 Mourão LC, Cataldo DM, Moutinho H, Fischer RG, Canabarro A (2014). Additional effects of homeopathy on chronic periodontitis: a 1-year follow-up randomized clinical trial. *Complement Ther Clin Pract*; **20**: 141–146.

A304 Thinesse-Mallwitz M, Maydannik V, Keller T, Klement P (2015). A homeopathic combination preparation in the treatment of feverish upper respiratory tract infections: An international randomized controlled trial. *Forsch Komplementmed*; **22**: 163–170.

A305 Chakraborty D, Das P, Dinda AK, Sengupta U, Chakraborty T, Sengupta J (2015). A comparative study of homoeopathic medicine – *Sulphur* with the Multidrug therapy in the treatment of paucibacillary leprosy. *Indian J Res Homoeopathy*; **9**: 158–166.

A306 Jong MC, Verwer C, van de Vijver L, Klement P, Burkart J, Baars E (2015). A randomized open comparative clinical trial on the effectiveness, safety and tolerability of a homeopathic medicinal product for the treatment of painful teething in children. *Altern Integr Med*; **4**: 1.

A317 van Haselen R, Thinesse-Mallwitz M, Maidannyk V, Buskin SL, Weber S, Keller T, Burkart J, Klement P (2016). The effectiveness and safety of a homeopathic medicinal product in pediatric upper respiratory tract infections with fever: A Randomized Controlled Trial. *Global Pediatric Health*; **3**: 1–11.

A318 Jong MC, Buskin SL, Ilyenko L, Kholodova I, Burkart J, Weber S, Keller T, Klement P (2016a). Effectiveness, safety and tolerability of a complex homeopathic medicinal product in the prevention of recurrent acute upper respiratory tract infections in children: a multicenter, open, comparative, randomized, controlled clinical trial. *Multidiscip Respir Med*; **11**: 19.

A319 Jong MC, Ilyenko L, Kholodova I, Verwer C, Burkart J, Weber S, Keller T, Klement P (2016b). 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*: 9539030.

MAIN-TEXT REFERENCES

[1] Swayne, J (2000). *International Dictionary of Homeopathy*, Churchill Livingstone, Edinburgh.

[2] Mathie RT, Legg LA, Clausen J, Davidson JRT, Lloyd SM, Ford I (2013). Systematic review and meta-analysis of randomised, placebo-controlled, trials of individualised homeopathic treatment: Study protocol. Version 1.0; 25 January. http://www.britishhomeopathic.org/wp-content/uploads/2013/05/Study_protocol_for_systematic_review.pdf

[3] Mathie RT, Legg LA, Clausen J, Davidson JRT, Lloyd SM, Ford I (2014). Systematic review and meta-analysis of randomised, placebo-controlled, trials of *non*-individualised homeopathic treatment: Study protocol. Version 1.0; 30 October 2014. http://www.britishhomeopathic.org/wp-content/uploads/2014/10/SR+MA-Protocol-Non-indiv-Hom-30-10-14-2-column++.pdf

[4] Mathie RT, Lloyd SM, Legg LA, et al (2014). Randomised placebo-controlled trials of individualised homeopathic treatment: systematic review and meta-analysis. *Systematic Reviews*; **3**: 142.

[5] Mathie RT (2015). Should doctors recommend homeopathy? The evidence is not a 'black or white' issue. *British Medical Journal*; **351**: h3735 (Published online, 14 July).

[6] Jacobs J, Jonas WB, Jimenez-Perez M, Crothers D (2003). Homeopathy for childhood diarrhea: combined results and metaanalysis from three randomized, controlled clinical trials. *Pediatric Infectious Disease Journal*; **22**: 229–234.

[7] Taylor MA, Reilly D, Llewellyn-Jones RH, McSharry C, Aitchison TC (2000). Randomised controlled trials of homoeopathy versus placebo in perennial allergic rhinitis with overview of four trial series. *British Medical Journal*; **321**: 471–476.

[8] Schneider B, Klein P, Weiser M (2005). Treatment of vertigo with a homeopathic complex remedy compared with usual treatments: a meta-analysis of clinical trials. *Arzneimittelforschung*; **55**: 23–29.

[9] Ernst E, Barnes J (1998). Are homoeopathic remedies effective for delayed-onset muscle soreness? – A systematic review of Placebo-controlled trials. *Perfusion (Nürnberg)*; **11**: 4–8.

[10] Ernst E (1999). Homeopathic prophylaxis of headaches and migraine? A systematic review. *Journal of Pain and Symptom Management*; **18**: 353–357.

[11] Ernst E (2011b). Homeopathy for insomnia and sleep-related disorders: A systematic review of randomised controlled trials. *Focus on Alternative and Complementary Therapies*; **16**: 195–199.

[12] Smith CA (2004). Homoeopathy for induction of labour. *Cochrane Database of Systematic Reviews*: CD003399.

[13] Mathie RT, Frye J, Fisher P (2015). Homeopathic Oscillococcinum® for preventing and treating influenza and influenza-like illness. *Cochrane Database of Systematic Reviews*; Issue 1: Article number CD001957.

[14] Long L, Ernst E (2001). Homeopathic remedies for the treatment of osteoarthritis: a systematic review. *British Homeopathic Journal*; **90**: 37–43.

[15] Kleijnen J, Knipschild P, ter Riet G (1991). Clinical trials of homoeopathy. *British Medical Journal*; **302**: 316–323.

[16] Boissel JP, Cucherat M, Haugh M, Gauthier E (1996). Critical literature review on the effectiveness of homoeopathy: overview of data from homoeopathic medicine trials. In: Homoeopathic Medicine Research Group, Report of the Commission of the European Communities, Directorate-General XII – Science, Research and Development, Directorate E – RTD Actions: Life Sciences and Technologies – Medical Research. Brussels, Belgium.

[17] Linde K, Clausius N, Ramirez G, et al (1997). Are the clinical effects of homoeopathy placebo effects? A meta-analysis of placebo-controlled trials. *Lancet*; **350**: 834–843.

[18] Cucherat M, Haugh MC, Gooch M, Boissel JP (2000). Evidence of clinical efficacy of homeopathy – A meta-analysis of clinical trials. *European Journal of Clinical Pharmacology*; **56**: 27–33.

[19] Linde K, Scholz M, Ramirez G, Clausius N, Melchart D, Jonas WB (1999). Impact of study quality on outcome in placebo-controlled trials of homeopathy. *Journal of Clinical Epidemiology*; **52**: 631–636.

[20] Shang A, Huwiler-Muntener K, Nartey L, et al (2005). Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. *Lancet*; **366**: 726–732.

[21] Lüdtke R, Rutten ALB (2008). The conclusions on the effectiveness of homeopathy highly depend on the set of analyzed trials. *Journal of Clinical Epidemiology*; **61**: 1197–1204.

[22] Ernst E (1999). Classical homeopathy versus conventional treatments: a systematic review. *Perfusion (Nürnberg)*; **12**: 13–15.

[23] Mathie RT (2003). The research evidence base for homeopathy: a fresh assessment of the literature. *Homeopathy*; **92**: 84–91.

[24] Sedgwick P (2013). Equivalence trials. *British Medical Journal*; **346**: f184.

[25] Akobeng AK (2008). Assessing the validity of clinical trials. *Journal of Pediatric Gastroenterology and Nutrition*; **47**: 277–282.

[26] Jüni P, Altman DG, Egger M (2001). Assessing the quality of controlled clinical trials. *British Medical Journal*; **323**: 42–46.

[27] Jadad AR, Moore RA, Carroll D, et al (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*; **17**: 1–12.

[28] Higgins JPT, Altman DG (2011). Chapter 8: 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.

[29] Bell IR (2003). Evidence-based homeopathy: Empirical questions and methodological considerations for homeopathic clinical research. *American Journal of Homeopathic Medicine*; **96**: 17–31.

[30] Sedgwick P (2012). External and internal validity in clinical trials. *British Medical Journal*; **344**: e1004.

[31] Bornhöft G, Maxion-Bergemann S, Wolf U, et al (2006). Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. *BMC Medical Research Methodology*; **6**: 56.

[32] Thorpe KE, Zwarenstein M, Oxman AD, et al (2009). A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology*; **62**: 464–475.

[33] Sense About Science (2013). Peer review. http://www.senseaboutscience.org/pages/peer-review.html [Accessed 13.04.16].

[34] Mathie RT, Hacke D, Clausen J, Nicolai T, Riley DS, Fisher P (2013). Randomised controlled trials of homeopathy in humans: characterising the research journal literature for systematic review. *Homeopathy*; **102**: 3–24.

[35] Homeopathy Research Institute. Randomised controlled trials of homeopathy: examining the evidence. https://www.hri-research.org/hri-research/learning-more-from-existing-evidence/systematic-review-programme/

[36] Baker DG, Myers SP, Howden I, Brooks L (2003). The effects of homeopathic *Argentum nitricum* on test anxiety. *Complementary Therapies in Medicine*; **11**: 65–71.

[37] Hewitt CE, Torgerson DJ, Miles JNV (2007). Is there another way to take account of non-compliance in randomized controlled trials? *Can Med Assoc J*; **175**: 347–348.

[38] Higgins JPT, Deeks JJ, Altman DG (2011). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions; Version 5.1.0.* The Cochrane Collaboration.

[39] Sedgwick P (2013). What is a non-inferiority trial? *British Medical Journal*; **347**: f6853.

[40] Deeks JJ, Higgins JPT, Altman DG (2011). Chapter 9: 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.

[41] Chinn S (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*; **19**: 3127–3131.

Appendix 1: **Eligible (*Yes*) and ineligible (*No*) combinations of intervention and comparator**

		Intervention	
		NIHT	**NIHT + Other Intervention #1**
Comparator	**Other Intervention #1**	*Yes*	*Yes*
	Other Intervention #2	*Yes*	*No*
	Nothing	*Yes*	*No*

NIHT: Non-individualised homeopathic treatment