Journal of Clinical Epidemiology 155 (2023) 64–72

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Homeopathy can offer empirical insights on treatment effects in a null field

Matthew K. Sigurdson[a,b], Kristin L. Sainani[a], John P.A. Ioannidis[a,b,c,d,e,*]

[a]Department of Epidemiology and Population Health, Stanford University, Stanford, CA, USA
[b]Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA
[c]Department of Medicine, Stanford University, Stanford, CA, USA
[d]Department of Biomedical Data Science, Stanford University, Stanford, CA, USA
[e]Department of Statistics, Stanford University, Stanford, CA, USA

Accepted 26 January 2023; Published online 1 February 2023

Abstract

Objectives: A ''null field'' is a scientific field where there is nothing to discover and where observed associations are thus expected to simply reflect the magnitude of bias. We aimed to characterize a null field using a known example, homeopathy (a pseudoscientific medical approach based on using highly diluted substances), as a prototype.

Study Design and Setting: We identified 50 randomized placebo-controlled trials of homeopathy interventions from highly cited meta-analyses. The primary outcome variable was the observed effect size in the studies. Variables related to study quality or impact were also extracted.

Results: The mean effect size for homeopathy was 0.36 standard deviations (Hedges' g; 95% confidence interval: 0.21, 0.51) better than placebo, which corresponds to an odds ratio of 1.94 (95% CI: 1.69, 2.23) in favor of homeopathy. 80% of studies had positive effect sizes (favoring homeopathy). Effect size was significantly correlated with citation counts from journals in the directory of open-access journals and CiteWatch. We identified common statistical errors in 25 studies.

Conclusion: A null field like homeopathy can exhibit large effect sizes, high rates of favorable results, and high citation impact in the published scientific literature. Null fields may represent a useful negative control for the scientific process.

Keywords: Homeopathy; Bias; Null field; Treatment effects; Meta-Research; Replication crisis; Research integrity

1. Introduction

In the search for causal and treatment effects, different fields of research may vary on whether they have a lot, a little, or nothing of essence to discover and many, few, or no effective treatments to offer. A ''null field'' has been defined as a field of research where there is nothing genuine to be discovered and no genuinely effective treatment exists [1]. Effect sizes are all null in their true magnitude [1,2]. Therefore, any observed effects reflect simply the magnitude of the prevailing biases in their composite influence [1]. Null fields are worth studying as negative controls to assess the amount of bias present in the scientific process. They can be used to investigate the overall effectiveness of the scientific process in avoiding false positives and indicate possible ways that it could be improved, e.g., for mitigation and reduction of bias.

Homeopathy is a form of alternative medicine which is expected to be a ''null field'' given basic scientific principles. Its treatments typically contain no active ingredient, and the claimed mechanisms are inconsistent with well-known physical laws [3,4]. It has also been used in the past as an example of a field with no substantive content [5]. Any observed effects may stem from bias, such as the

There are no conflicts of interest in this submission.

Author contributions: Matthew K. Sigurdson: Conceptualization, Methodology, Software, Validation, Formal Analysis, Data Curation, Investigation, Writing - Original Draft, Writing - Review & Editing. Kristin L. Sainani: Conceptualization, Methodology, Writing - Review & Editing, Supervision. John P. A. Ioannidis: Conceptualization, Methodology, Writing - Review & Editing, Supervision.

Funding: The Meta-Research Innovation Center at Stanford (METRICS) has been funded by the Laura and John Arnold Foundation.

Ethical approval: none.

* Corresponding author. University of Stanford, Medical School Office Building, Room X306, Stanford, CA 94305, USA. Tel.: +1-650-725-5465; fax: +1-650-725-6247.

E-mail address: jioannid@stanford.edu (J.P.A. Ioannidis).

https://doi.org/10.1016/j.jclinepi.2023.01.010
0895-4356/

What is New?

Key findings

- In our analysis, studies evaluating homeopathy interventions found an average treatment effect of g = 0.36 standard deviations compared to control.

What this adds to what was known?

- The average observed effect of these homeopathy interventions reflects an empirical estimate for the average bias present and indicates the expected result when no effect exists. This may allow empirical calibration of effect sizes.

What is the implication and what should change now?

- Null fields, such as homeopathy, can be used as negative controls for the scientific process, by examining what effect sizes they publish and the characteristics of these studies.

placebo effect, but also other biases. It is ideal as a prototypical null field, since it effectively evaluates placebo treatments with no possible other effects or side effects and has a large body of scientific literature. As of June 2022, a PubMed search for *homeopathy OR homoeopathy* returned over 6,200 results.

We performed a descriptive analysis of homeopathy as a prototypical null field. Using a sample of 50 studies of homeopathic treatments, selected from the most highly cited meta-analyses in the field, we used the observed effect size distribution to estimate the typical strength of the bias. We also evaluated the characteristics of the field, focusing on those that could potentially be associated with null fields and investigated factors that may be correlated with reported effect sizes, *P*-values, and positive conclusions.

2. Methods

2.1. Search strategies and eligibility criteria

We used a two-phase search strategy in which homeopathy meta-analyses were identified first, and individual studies were then selected from those included in the meta-analyses (Fig. 1). The meta-analyses were retrieved from PubMed using the search term *homeopathy OR homoeopathy* combined with either *meta-analysis[Publication Type]* or *\"Cochrane Database Syst Rev\"[Journal]*. The results of both searches were retrieved in November 2020 using the "RISmed" package in R [6], which retrieves data such as publication year and total citation counts from PubMed, and were ordered based on citations per year in PubMed since publication. Meta-analyses were included if they performed a literature search and meta-analysis for a relationship involving homeopathy. A total of 50 unique studies were retrieved from the meta-analyses, with a maximum of 10 studies from each meta-analysis in order to reduce any undue influences from any single meta-analysis. Within meta-analyses with more than 10 studies, the 10 studies were selected randomly using a computer-based pseudorandom number generator. All papers in PubMed written in English that could be obtained were eligible. Since the most highly cited meta-analyses were used first, this strategy is largely expected to retrieve influential studies in the field.

2.2. Data extraction and analysis

Our primary outcome variable is effect size. To allow comparisons between different studies, effect sizes were retrieved as standardized mean deviation (SMD), specifically Hedges' g, calculated as the ratio of the effect size (change in the treatment group compared to change in the control group, or final measurement in the treatment group compared to final measurement in the control group) to the baseline standard deviation multiplied by a correction factor. Binary outcomes were retrieved as odds ratios. SMD and odds ratio values were interchanged using an existing method [7]. Other measurements were recalculated using data from the studies, such as differences in proportions, which were recalculated as odds ratios. We corrected errors that affected the calculation of the effect size; for example, if authors calculated the effect size by comparing the homeopathy group to itself at baseline (a within-group comparison), we corrected this to an effect size comparing the homeopathy group to the control group (a between-group comparison). If necessary, the sign of the measurement was reversed or the reciprocal was taken in order to allow for comparisons with other studies. The specific variable retrieved was the primary outcome variable of the paper, if one was specified, or a substitute was chosen based on the predetermined protocol (summarized in the supplementary information). If the published study did not have the information necessary to calculate a valid effect size with an associated measurement of variation, then a reasonable estimate or a different outcome variable was used instead, following the predetermined protocol.

We also extracted data on other variables that are associated with study quality or impact, such as citation metrics and the presence of errors. The full dataset, including the extracted data needed to reproduce the effect size calculations, is included in a separate file (see the dataset provided in the supplement), and exact variable definitions are reported in the data dictionary. In addition to PubMed, author information was obtained from Scopus and journal impact data from Journal Citation Reports. To retrieve information from figures, WebPlotDigitizer [8] was used. Information was extracted automatically where possible, principally with "RISmed" or "rentrez" [6,9], and otherwise was recorded manually by MKS, with cases of ambiguity resolved by consulting with KS. If a study was missing information or could be interpreted in different ways, variables were recorded as ‘‘unclear’’. Such entries were treated as missing data for continuous variables but negative for binary variables, following precedent [10,11]. In addition, we generally attempted to extend the benefit of the doubt to the study authors where possible. For example, in the identification of errors, the absence of a multiple testing correction was defined conservatively to include only cases where the need for a correction was unambiguous.

In total, we recorded 103 variables; this count includes basic information (such as PubMed ID) and sets of highly correlated variables that address the same topic (such as standard deviation and *P*-value, or number of citations received and number of citations received per year). Indicators showing that a journal may be considered ‘‘questionable’’ were the absence of an impact factor, or indexing status in MEDLINE, Web of Science, Directory of Open Access Journals (DOAJ) [12], or The Wikipedia CiteWatch [13]. As an exploratory analysis, we tested for correlations for 77 variables (numeric or binary variables for which a correlation measurement would be meaningful, e.g., excluding PubMed ID) with effect size and two other variables: the magnitude of the *P*-value and whether the authors included a positive statement about the intervention's effectiveness as part of their main conclusion in the abstract (or the discussion, if the abstract had no conclusion-like statement). Correlations were measured as Spearman's correlation coefficients. Due to the large number of explored variables, correlations significant at $P < 0.05$ only represent tentative signals, and we also performed correction for multiple comparisons using the Holm−Bonferroni method. All analyses were performed in R version 4.1.1.

3. Results

50 homeopathy intervention studies were selected from 13 meta-analyses. The estimated mean effect size, measured as Hedges' g, was 0.36 (95% confidence interval [CI]: 0.21, 0.51) (Fig. 2), indicating a sufficient degree of bias to produce an overall result which is significantly different from placebo. This corresponds to an odds ratio of 1.94 (95% CI: 1.69, 2.23) in favor of homeopathy. In contexts where a study measures odds ratios less than 1, e.g., investigations of protective effects, the corresponding odds ratio is the reciprocal, 0.52 (95% CI: 0.45, 0.59). The weighted mean was g = 0.29. Overall, the direction of the effect favored homeopathy for 40 studies and favored

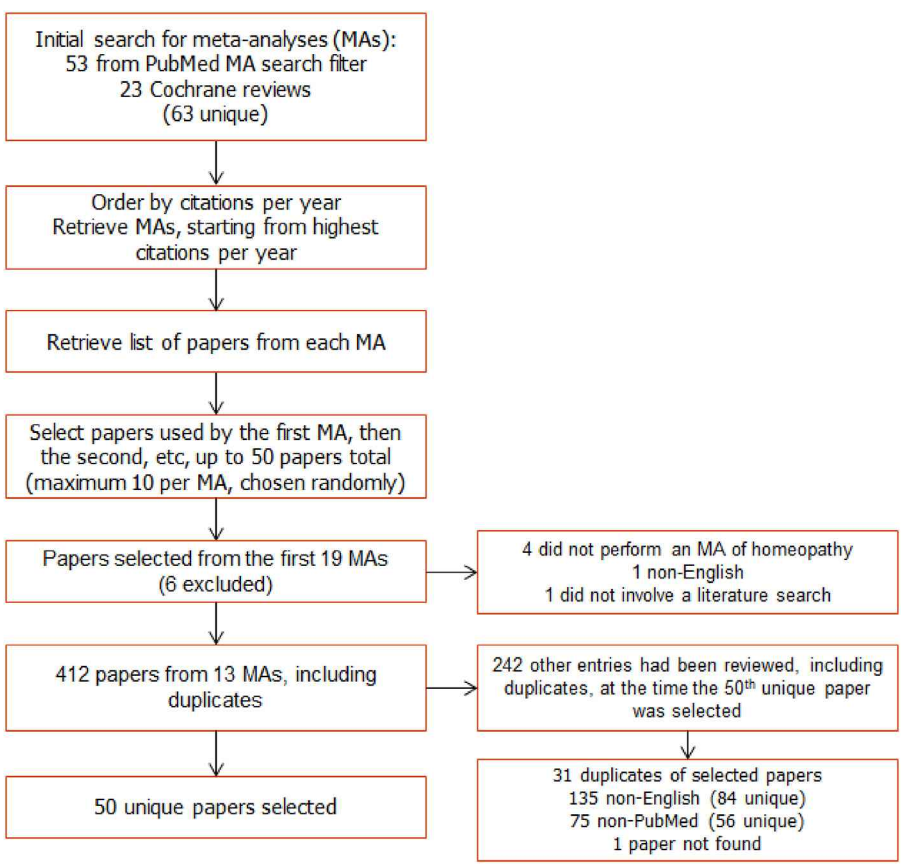

Fig. 1. Flow diagram of the search strategy. If multiple exclusion criteria applied to a study, only one was counted here.

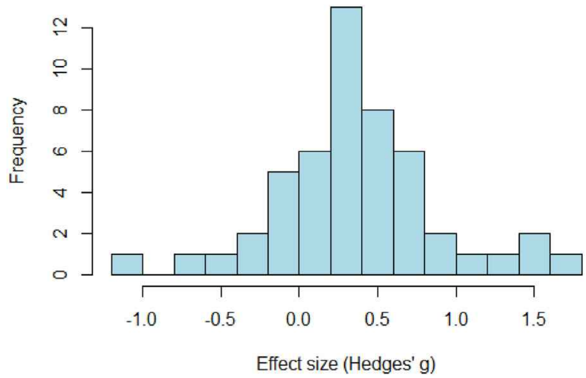

Fig. 2. Effect size distribution in homeopathy (n = 50), measured as Hedges' g. The mean of the distribution is 0.36 (95% CI: 0.21, 0.51), which corresponds to an odds ratio of 1.94 (95% CI: 1.69, 2.23) or its reciprocal 0.52 (95% CI: 0.45, 0.59).

control for 9 studies, and one study had a difference of exactly 0. There were 19 studies statistically significant at the 0.05 level, 17 that favored treatment, and 2 that favored control. Neither of the studies that significantly favored control reported their results as such: one calculated their outcome differently due to a math error, and the other analysis was nonsignificant because it did not optimally account for baseline differences. The distribution of treatment effects was slightly right-skewed with a median Hedges' g of 0.33 (interquartile range [IQR] = 0.53). The Pearson median skewness was 0.17, in contrast to a plausible alternative hypothesis, where large positive effect sizes might be heavily over-represented in null fields.

Descriptive and analytic characteristics of the 50 studies are shown in Table 1 and Table 2. All 50 studies were randomized controlled trials. The interventions and outcomes were diverse, and for 21 studies the intervention tested was an "individualized homeopathic remedy", reflecting the claim by homeopaths that their treatments must be chosen separately for each patient. Among the studies that tested a specific treatment, few of them tested the same one, and those that did were often at very different doses. Among the outcomes, common themes were allergy, pain, and symptom relief.

The studies were published in 31 unique journals. Among the 40 studies published in journals with an impact factor, most are relatively low-impact, but several are quite prestigious, including the *Lancet* and *BMJ*. The most common journals were *Homeopathy* (formerly *The British Homoeopathic Journal*) (10 studies), *BMJ* (3 studies), *British Journal of Clinical Pharmacology* (3 studies), *Journal of Alternative and Complementary Medicine* (3 studies), and *Lancet* (3 studies). Many of the authors are relatively prolific publishers, based on the number of publications over their career, and are also highly cited based on h-index and total number of citations received. The number of citations received by the individual studies was relatively more modest, although only references in PubMed were counted. Several of the journals where the studies were published, for both the individual studies and those that cited them, are classified as potentially questionable based on several different metrics (Table 2).

Many studies were missing key elements that are usually viewed as essential, such as power calculations and discussion of limitations. Most outcome variables were subjective, usually based on judgements from the patient with a minority coming from the treatment provider. There were 25 studies for which an error was discovered in one of the three following categories: multiple testing, presenting a within-group change as a between-group difference (as described above) and math errors. A multiple testing error occurs if there were multiple statistical comparisons performed that had a systematic relationship, such as the same outcome compared at multiple timepoints or across multiple body parts, and no correction for multiple testing was used. A within-group error occurs if authors calculate the effect size incorrectly by comparing the homeopathy group to itself at baseline instead of comparing it directly to the control group (for example, inferring that the treatment group did better if a measurement changed significantly in the treatment group but not in the control group). In four studies, there were errors discovered in two categories, for a total of 29 errors observed. Other categories of errors were not recorded.

In 21 studies (42%), the main conclusion did not report a valid main effect, meaning that it either presented a comparison that did not address whether the treatment works, or it did not accurately state whether the comparison they had reported was statistically significant or nonsignificant. 34 studies (68%) included a component in their main conclusion which was positive about the effectiveness of homeopathy, 17 of which were not justified by the primary analysis. This becomes 18/35 (51%) when also including the study which had no positive or negative component when the primary analysis favored control, which is also a change favorable to homeopathy. The two most common sources for the homeopathy-favoring statement were within-group comparisons (5 studies, 28%) and nonsignificant results, such as $P > \alpha$ interpreted as "trends" (4 studies, 22%). Other sources included errors, focusing on a subgroup, and descriptions of other variables.

In the correlation analysis (Table 3), the effect size and the other two variables tested, P-value and presence of a positive statement in the conclusion, all correlated with each other after correction for multiple testing. Three other correlations met the same standard: the effect size correlated with the number of citations received from high-quality journals indicated by DOAJ indexing and from potentially questionable journals indicated by CiteWatch indexing, and the presence of a positive statement in the conclusion correlated with the presence of a demographics table which described the characteristics of the sample. Several other correlations were only significant before multiplicity correction and they mostly pertained to the number and properties of citations and other impacts and the number of authors as correlates of better-looking results.

Table 1. Characteristics of studies in homeopathy.

Variable	Measurement (n = 50)	Range
Publication year	2000 (12.5)	1978–2012
Sample size	56 (55.75)	16–462
Placebo-controlled	50 (100)	–
Clinical trial	50 (100)	–
Randomized	50 (100)	–
Randomized adequately	16 (32)	–
Blinding		
Patient (“single”)	42 (84)	–
Patient and provider (“double”)	35 (70)	–
Analyst	15 (30)	–
Primary outcome variable identified	27 (54)	–
Number of outcome variables	6.5 (6.5)	2–46
Intention to treat (ITT)	9 (18)	–
Modified ITT	19 (38)	–
Number of authors	5 (2)	1–15
First author		
Number of publications	38 (85.5)	1–643
Publications/year	2 (1.55)	0.12–22.17
Total citations received	450.5 (1,100.75)	8–26,302
h-index	10 (15.25)	1–93
Last author		
Number of publications	59 (152.5)	1–643
Publications/year	2.2 (3.1)	0.28–22.17
Total citations received	1,306 (3,340)	8–26,302
h-index	19 (25.75)	1–93
Open access	13 (26)	–
Demographics table reported	37 (74)	–
Number of references cited	17 (18.25)	1–63
PubMed citations received		
Overall	7.5 (9.5)	0–67
Reviews	1.5 (2.75)	0–26
Meta-analyses	1 (1)[a]	0–3
PubMed citations/year received		
Overall	0.35 (0.36)	0–3.53
Reviews	0.06 (0.14)	0–1.37
Meta-analyses	0.03 (0.05)[a]	0–0.16

Continuous variables are presented as median (IQR) and have an associated range in the right-hand column; categorical variables are n (%) and do not have an associated range. Exact variable definitions are reported in the data dictionary.

[a] Many of the studies n have 0 for the number of citing meta-analyses, even though the search strategy required every study in the sample to be included in at least one meta-analysis, because many meta-analyses do not include the studies they analyzed in their formal list of references.

4. Discussion

We have assessed effect sizes and their correlates in homeopathy as a prototypical example of a null field. We found that studies in homeopathy have an average effect size of 0.36 standard deviations, which is quite substantial. Since homeopathy is known to be a null field, this estimate does not indicate that there is a true effect; instead, it is the average impact of the bias present in the field. Because these are all randomized, placebo-controlled trials, a theoretically rigorous study design, typical effect sizes due to bias in other, theoretically less rigorous designs, may be even greater in this field and other null fields.

Previous analyses have found that studies in homeopathy have similar [14] or better [11] “quality” features than studies in conventional fields, based on a small number of common evaluation criteria (e.g., randomization, blinding, intention-to-treat analysis), although both groups scored

Table 2. Potential indicators of study reliability.

Variable	Measurement (n = 50)	Range
Subjective primary outcome variable	44 (88)	–
Source of subjectivity		
Patient	34 (68)	–
Provider	8 (16)	–
Both	2 (4)	–
Preregistration	2 (4)	–
Journal quality indicators		
Impact factor published	40 (80)	–
Impact factor of journal (n = 40)[a]	2.55 (3.54)	0.49–59.21
MEDLINE indexing	36 (72)	–
Web of Science (WOS) indexing	40 (80)	–
Directory of Open Access Journals (DOAJ) indexing (n = 8)[b]	4 (50)	–
CiteWatch (n = 45)[c]	0 (0)	–
Journal quality indicators of journals citing the underlying studies (n = 47 with citations in PubMed)[d]		
Proportion with impact factor published	0.89 (0.17)	0.56–1
Average impact factor[a]	5.06 (6.67)	0.91–54.52
Proportion with MEDLINE indexing	0.62 (0.32)	0–1
Proportion with WOS indexing	0.88 (0.24)	0.33–1
Proportion with DOAJ indexing[b]	0.83 (0.5)	0–1
Proportion CiteWatch-flagged	0.09 (0.23)	0–1
Subgroup analyses included	22 (44)	–
Pre-study power calculation	18 (36)	–
Statistical methods section	40 (80)	–
Adjustment for confounding considered	17 (34)	–
Error(s) identified		
Total	25 (50)	–
Multiple testing correction missing where needed	18 (36)	–
Within-group comparison misinterpreted as		
Between-group comparison	9 (18)	–
Math error	2 (4)	–
Limitations discussed (min. 1 paragraph)	20 (40)	–
Valid main effect reported in main conclusion	29 (58)	–
Positive statement about intervention's effectiveness reported in main conclusion	34 (68)	–
Homeopathy-favoring statement is inconsistent with primary analysis (n = 35)[e]	18 (51)	–
Source of homeopathy-favoring statement (n = 18)		
Within-group comparison	5 (28)	–
Nonsignificant result	4 (22)	–
Other	9 (50)	–
Continued research recommended	31 (62)	–
Funding statement included	30 (60)	–
Government funding disclosed (n = 30)	11 (36.7)	–
Corporate funding disclosed (n = 30)	9 (30)	–
Nonprofit funding disclosed (n = 30)	13 (43.3)	–
Conflict of interest statement included	6 (12)	–

Continuous variables are presented as median (IQR) and have an associated range in the right-hand column; categorical variables are n (%) and do not have an associated range. Exact variable definitions are reported in the data dictionary. Indicators of reliability already shown in Table 1 (e.g., sample size, blinding) are not repeated here

[a] The second impact factor measurement is expected to be higher than the first one, because it is the median of the mean values instead of solely being the median.

[b] Only open-access journals were evaluated. For the proportion measurement, n = 45 instead of 47. DOAJ: Directory of Open Access Journals.

poorly. Proposed reasons included the relative ease of performing double-blinding and allocation concealment in homeopathic studies [14] and a higher degree of heterogeneity in the trials of conventional medical interventions due to the presence of genuine treatment effects [11]. One of the analyses [14] noted that the quality was not in fact comparable despite the similar quality scores, citing other criteria, such as the use of clear outcome measures. Selected studies in homeopathy assessments may meet standards such as double-blinding, but not less tangible indicators. Alternatively, similarities between null and non-null fields could be produced by the null field ‘‘adapting’’ through selective pressure to display the same characteristics of design and reporting as regular fields, because meeting specific standards improves the chance of publication. In this context, the null field could evolve to be even better at specifically displaying the features that are most often used as selection criteria by journals, or could develop ways to subvert them.

Unsurprisingly, effect sizes were negatively correlated with *P*-values and positively correlated with positive statements in the study conclusions. The correlation of effect size with the number of citations from DOAJ-indexed and CiteWatch-flagged journals suggests that more prominent effect sizes and more promising results attract citations both by legitimate and questionable journals. While none of the studies in the sample were themselves published in CiteWatch-flagged journals, likely related to the quality threshold imposed by our search algorithm, the studies citing them were not under the same constraints. With regards to the demographics table, the correlation with positive conclusions may reflect that ‘‘positive’’ trials are given more space and elaboration for presenting their work.

Null fields, such as homeopathy, can be thought of as negative controls for science, indicating the degree to which effects can be observed when no actual effect is present. This approach can be used to estimate the overall magnitude of bias and evaluate the effectiveness of methods for mitigating bias. By investigating how the process of science can fail, one can attempt to identify similar cases or prevent them from happening again in the future.

It is often assumed that when a treatment or risk factor has no effect, the observed results should be null (e.g., an odds ratio of 1). However, this is a misconception. Observed published results are actually expected to reflect the level of bias present in the study, an important distinction which is not accounted for in standard null hypothesis significance testing (NHST). Whenever possible, one should be making ‘‘calibrated’’ comparisons to the expected level of bias rather than to the number that represents a null effect. A homeopathy trial with a g = 0.36, therefore, shows a typical effect, that may be the average transformation of the null. Similar calibration considerations have been proposed and demonstrated empirically for observational studies [15].

Methods to address the effects of biases [16,17] are not widely used in epidemiology [18,19], and many only allow one type of bias to be addressed at a time. Estimating the impact of the total bias may be particularly valuable because that has the most direct impact on the final result, and it is unlikely that all major sources of bias acting on a study could be measured or even enumerated.

The magnitude of overall bias found in the current study is consistent with measurements of the average effects of individual biases. One analysis of 1,973 randomized controlled trials found that on average, the effect size was exaggerated by 11% with inadequate or unclear random sequence generation, 7% with inadequate or unclear allocation concealment, and 13% with absent or unclear double-blinding. The impact was greatest in trials with subjective outcomes [10]. Different study designs have also been found to be biased in favor of detecting significant results. In one investigation, instead of the presumed 5% false positive rate, the false positive rate was 50% in cohort studies and 72% in case-control studies [15]. The impacts of specific analytic choices have also been studied [20].

Meta-analysis may offer an opportunity to reduce some biases. However, meta-analyses are not necessarily sufficient, at least in the case of homeopathy (‘‘garbage in, garbage out’’ is a well-known criticism) [21]. Furthermore, meta-analysis increases the chances that summary results will become more statistically significant when it combines multiple biased studies. In the case of homeopathy, we found that out of the 7 meta-analyses in our sample that reported an overall effect size, 6 favored homeopathy, with 4 reported as statistically significant (Table S1).

We should caution that our study is primarily descriptive, and does not address causal relationships. The current analysis does not perform direct comparisons with the characteristics of other fields, and other null fields may not show the same characteristics. The degree and types of bias can be considerably different across fields [22]. This may be affected by treatment characteristics, such as treatments which are more difficult to blind, or by the culture of the field itself. For instance, another possible indicator of biased studies is that effect sizes are implausibly inflated [23], so a null field in which this is common would have a heavily right-skewed null distribution. Null fields

[c] Only DOIs were checked against the CiteWatch list; the remaining studies did not have associated DOIs.

[d] For each study, citations received were evaluated for the characteristics of the journals where the citations were published. For each of the described journal characteristics, a proportion was calculated for each of the 47 underlying studies, and the reported values are the median, IQR, and range for those proportions. For average impact factor, the reported values are the median, IQR, and range for the average.

[e] In addition to the 34 studies with positive statements in their main conclusion, this includes one study which found nonsignificant results instead of results significantly favoring placebo, since this is also a change favorable to homeopathy. There were two such studies in total (see text), but the other was already included in the 34 with positive statements in their main conclusion.

Table 3. Variables associated with effect size, significance, and positive author descriptions (Spearman's correlation).

Variable	Correlation with effect size	Correlation with *P*-value	Correlation with positive statement in conclusion
Effect size	–	−0.66*	0.63*
P-value	−0.66*	–	−0.58*
Positive statement in conclusion	0.63*	−0.58*	–
Number of DOAJ-indexed citations	0.49*	−0.35	0.32
Number of CiteWatch-flagged citations	0.47*	−0.45	0.29
PubMed references received/year	0.41	−0.41	0.36
Proportion of CiteWatch-flagged citations	0.41	−0.37	–
Number of citations with published IF	0.39	−0.36	–
Number of WOS-indexed citations	0.39	−0.34	–
PubMed references received	0.35	−0.37	–
Demographics table included	0.32	–	0.57*
Number of unique citing journals	0.31	−0.35	–
PubMed review references received/year	–	−0.30	0.39
Number of positive subgroup analyses	–	−0.29	0.31
Number of references cited	–	−0.29	0.34
PubMed review references received	–	−0.29	0.30
PubMed citations from the same journal	0.29	–	–
Number of authors	0.28	–	–
Number of dropouts and missing data	–	−0.32	–
Further research recommended	–	–	0.35

DOAJ, directory of open access journals; WOS, web of science; IF, impact factor.

Only correlations that reached nominal significance ($P < 0.05$) are included. Values with asterisks are still significant after Holm−Bonferroni multiple testing correction.

could also be the product of different circumstances, which could lead to different characteristics. In addition, the patterns of bias prevalent within a field may change over time.

Studying multiple biased fields may offer complementary insights. Furthermore, heterogeneity is possible even within our sample, especially given the wide variety of interventions and outcomes observed. Nevertheless, on average, the absence of treatment effects would be expected to lead to less heterogeneity. In addition, the treatments are unified by the underlying principles in use. Homeopathy in particular is easier to justify in this regard, because every treatment is effectively identical, and also identical to a placebo, regardless of what exposure is allegedly being tested. Our analysis only included placebo-controlled trials that in principle would remove the placebo effects that homeopathy may achieve. However, biases such as poor masking or inadequate allocation concealment may have damaged the integrity of the placebo design in some trials. Such biases simply add to the overall bias.

Finally, a number of fields of research may have only a small proportion of their research targets being genuinely non-null. If so, their effect distribution and characteristics may approximate null fields. For example, some subfields of nutrition research may be approximately null [24,25]. Apparent effects on the edge of detectability are one of the hallmarks of pathological science [26,27], and other fields of alternative medicine may also be null or approximate null fields. Methods to identify and study null and approximately null fields may help reduce research waste and improve the allocation of scientific resources.

Acknowledgments

We are grateful to John Borghi for advice on the study protocol.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2023.01.010.

References

[1] Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2:e124.

[2] Sijtsma K. Playing with data–or how to discourage questionable research practices and stimulate researchers to do things right. Psychometrika 2016;81:1−15.

[3] Grimes DR. Proposed mechanisms for homeopathy are physically impossible. Focus Altern Complement Therapies 2012;17:149−55.

[4] Cukaci C, Freissmuth M, Mann C, Marti J, Sperl V. Against all odds-the persistent popularity of homeopathy. Wien Klin Wochenschr 2020;132:232−42.

[5] Reisman S, Balboul M, Jones T. P-curve accurately rejects evidence for homeopathic ultramolecular dilutions. PeerJ 2019;7:e6318.

[6] Kovalchik S. RISmed: download content from NCBI databases. R package version 2.3.0. 2021. Available at: https://CRAN.R-project.org/package=RISmed. Accessed January 11, 2020.

[7] Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med 2000;19:3127−31.

[8] Rohatgi A. WebPlotDigitizer v. 4.5. 2021. Available at: www.automeris.io/WebPlotDigitizer. Accessed January 2, 2021.

[9] Winter DJ. rentrez: an R package for the NCBI eUtils API. R J 2017;9:520−6.

[10] Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med 2012;157:429−38.

[11] Shang A, Huwiler-Muntener K, Nartey L, Juni P, Dorig S, Sterne JA, et al. Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. Lancet 2005;366:726−32.

[12] Lund University Libraries. Directory of Open Access Journals (DOAJ). United Kingdom: Infrastructure Services for Open Access; 2003.

[13] Wikipedia. (2021) The Wikipedia CiteWatch.www.en.wikipedia.org/wiki/Wikipedia:WikiProject_Academic_Journals/Journals_cited_by_Wikipedia/Questionable1. Accessed October 3, 2021. Described at www.en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/2019-2003-2031/In_focus under its former name, Wikipedia SourceWatch.

[14] Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV, et al. Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials. Lancet 1997;350:834−43.

[15] Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Stat Med 2014;33:209−18.

[16] Greenland S. Multiple-bias modelling for analysis of observational data. J R Stat Soc A Stat 2005;168:267−91.

[17] Timothy L, Lash MPF, Fink Aliza K. In: Applying Quantitative Bias Analysis to Epidemiologic Data. 1 ed. New York, NY: Springer; 2009.

[18] Innes GK, Bhondoekhan F, Lau B, Gross AL, Ng DK, Abraham AG. The measurement error elephant in the room: challenges and solutions to measurement error in epidemiology. Epidemiol Rev 2022;43:94−105.

[19] Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol 2014;43:1969−85.

[20] Madigan D, Ryan PB, Schuemie M. Does design matter? Systematic evaluation of the impact of analytical choices on effect estimates in observational studies. Ther Adv Drug Saf 2013;4:53−62.

[21] Ioannidis JP, Lau J. Can quality of clinical trials and meta-analyses be quantified? Lancet 1998;352:590−1.

[22] Fanelli D, Costas R, Ioannidis JP. Meta-assessment of bias in science. Proc Natl Acad Sci USA 2017;114:3714−9.

[23] Ioannidis JP. Exposure-wide epidemiology: revisiting Bradford Hill. Stat Med 2016;35:1749−62.

[24] Ioannidis JPA. The challenge of reforming nutritional epidemiologic research. JAMA 2018;320:969−70.

[25] Brown AW, Ioannidis JP, Cope MB, Bier DM, Allison DB. Unscientific beliefs about scientific topics in nutrition. Adv Nutr 2014;5:563−5.

[26] Langmuir I. Pathological science. Res Technol Manage 1989;32:11−7.

[27] Elton DC, Spencer PD. Pathological water science − four examples and what they have in common. In: Gadomski A, editor. Water in Biomechanical and Related Systems. Cham: Springer International Publishing; 2021:155−69.