The 2005 Lancet review proved superior quality of homeopathy trials: what's next?

Alexander Leonard Benedictus Rutten
Committee for Methods and Validation VHAN, Breda, Netherlands

ABSTRACT

Several former reviews showed positive effects of homeopathy, but in 2005 The Lancet published a review which claimed that homeopathy is a placebo effect. This review was criticised for not revealing essential information. A reconstruction of post-publication data challenges the negative conclusion. The only conclusion that was rectified by the methodology of the 2005 review was that the quality of homeopathy trials, and especially of smaller trials, is better than quality of conventional trials. The comparison of the effect of 110 homeopathy trials with 110 matched conventional trials was flawed by selection bias, different publication bias, different quality, and different safety. Nevertheless, there is no significant difference of effects between both methods. Discussions about proof for homeopathy are in fact discussions about science. The prior assumption that homeopathy cannot work pervades all aspects of this discussion and is not properly evaluated in the introduction of most analyses.

Keywords: homeopathy, review; Lancet 2005

Introduction

In 1991 Kleijnen et al published a meta-analysis of 105 homeopathy trials concluding: “The amount of positive evidence even among the best studies came as a surprise to us. Based on this evidence we would readily accept that homeopathy can be efficacious, if only the mechanism of action were more plausible... the evidence presented in this review would probably be sufficient for establishing homeopathy as a regular treatment for certain indications” [1]. In 1997 Linde et al's conclusion, based on 89 trials, was: “The results were not compatible with the hypothesis that the effects of homeopathy are completely due to placebo” [2]. In 1998, Linde acknowledged that studies with better quality tended to yield less positive results [3]. Still, Vandenbroucke stated in 2001: “A reflection of the scientific behaviour of adherents of conventional medicine toward one form of alternative medicine – homeopathy – teaches us that physicians do reject seemingly solid evidence because it is not compatible with theory” [4]. Apparently meta-analyses do not end the discussion about the efficacy of homeopathy due to other factors. After analysing a large number of reviews from the Cochrane database Ezzo concluded: “The number of reviews indicating that the modern biomedical interventions show either no effect or insufficient evidence is surprisingly high. Interrater disagreements suggest a surprising degree of subjective interpretation involved in systematic reviews” [5].

In 2005 an analysis of homeopathy by Shang et al, a team led by Egger, concluded that homeopathy is a placebo-effect [6]. This conclusion, however, was based on an undisclosed set of eight trials despite requests from the steering committee for this analysis four months before publication. The subsequently provided data showed that the conclusion highly depended on subjective choices and one indication (muscle soreness) and could well be a false-negative result [7]. The methodology in this analysis departed from the original hypotheses about quality and made comparisons of effect based on incomparable subgroups [8].
The predefined hypotheses in the introduction of Shang’s analysis were: “Bias in conduct and reporting of trials is a possible explanation for positive findings of placebo-controlled trials of both homeopathy and allopathy (conventional medicine)”; and: “These biases are more likely to affect small than large studies; the smaller a study, the larger the treatment effect necessary for the results to be statistically significant, whereas large studies are more likely to be of high methodological quality and published even if their results are negative”. The reporting about these two hypotheses was incomplete. The overall quality of homeopathy in this comparative analysis was better: Homeopathy had 21 (19%) good quality trials, conventional medicine 9 (8%). The quality of smaller trials was not mentioned.

The discussion about homeopathy seems to focus on quality bias. Egger suggested that quality bias requires regression analysis evaluating funnel plot asymmetry instead of random effects analysis, as applied in former analyses of homeopathy [9]. In this paper the function of regression analysis and quality of homeopathy in smaller trials is explored. What conclusions can be drawn from a comparison of quality regarding comparison of efficacy of homeopathy and conventional medicine? And how should this discussion go on?

Quality bias and regression analysis

In 1997 Egger proposed to evaluate the asymmetry of the funnel plot to discover bias [10]. The idea is represented in figure 1 in a simplified way. Egger’s assumption is that smaller trials are more biased, e.g. by insufficient concealment of placebo allocation. If doctors can identify the patients who receive placebo their assessment of results is influenced by this knowledge, producing exaggerated effects. The lower dot in figure 1 represents the average smaller trial. If larger trials are less biased the effect will be smaller, the upper dot in figure 1 represents the average larger trial. If you follow the regression line between smaller and larger trials you will find the result with no bias in the top, representing the largest possible trial.

Figure 1: The influence of quality bias on effect. Upper dot represents the average large trial, the lower dot the average small trial. The effect of the small trial is over-estimated by quality bias. The real effect is obtained by extrapolation towards the largest possible trial.

The assumption that smaller trials are more biased is criticised by Vandenbroucke: “What if the larger trials are those that were stopped judiciously at the right moment or underwent some data-analytic ‘massage’?” [11]. Egger himself added another caveat to the use of regression in this respect [9]. Trials in smaller populations can show stronger effects for a valid reason, if patients are better selected. Such selection could concern patients with stronger disease or better selection of patients susceptible to the medicine by genetic tests. Better selection of patients to obtain larger effects is essential in homeopathy. This selection is not based on genetic tests but on bayesian application of practice experience [12]. If, say, patients that respond well to the homeopathic medicine Lachesis are more frequently loquacious than the remainder of the population, the fact that a new patient is loquacious will increase the chance that Lachesis will work. Selection of susceptible patients therefore leads to larger effects in smaller trials than in larger trials; and vice versa. What can be done about this problem? Can you provide a comparison of quality regarding comparison of efficacy of homeopathy and conventional medicine? And how should this discussion go on?
patients will cause smaller research populations. This aspect is shown in figure 2: the stronger effect in a smaller population is real and not caused by bias. If there is no bias we should not use extrapolation towards the largest trials. It is therefore essential to discern between smaller trials with and without quality bias.

Figure 2: The influence of better selection of patients on the real effect. The upper dot represents the effect in an unselected (therefore larger) group of patients. After selecting sensitive patients, resulting in a smaller sample, effect increases. In this case the optimal effect is a real effect, not over-estimated by bias.

Comparison of quality in smaller trials

Comparing homeopathy and conventional medicine by regression meta-analysis is only possible if quality, and especially quality of smaller trials, is the same in both methods. The problem is that trials should also be comparable regarding indications. Shang’s analysis started as a comparison of 110 homeopathy trials with 110 conventional trials matched by indications. This matching by indication gets lost by selecting subgroups on other criteria, like quality and sample size. The cut-off value for ‘smaller trial’ is rather arbitrary, one could choose values like median sample size or n<100. The degrees of matching for these cut-off values were:

| Trials comparison            | Degree of matching |
|------------------------------|--------------------|
| 110 homeopathy - 110 conventional | 100%               |
| Trials with n<100            | 82%                |
| Trials with n<66 (median)    | 65%                |

The most valid cut-off value for ‘smaller trials’ seems to be n<100. There were 14 homeopathy studies of higher quality out of 83 trials (16.9%) with n<100. There were two conventional studies of higher quality out of 78 trials (2.6%) with sample size <100. The hypothesis that especially low quality small studies are responsible for the positive findings in homeopathy is falsified (p=0.003, Fisher exact probability test). There is statistically significant difference in quality of smaller studies in favour of homeopathy.

Meta-regression analysis in homeopathy

If we estimate the effect of all 21 good quality trials as selected by Shang by random effects meta-analysis homeopathy is effective, OR=0.76 (95% CI: 0.59-0.99). Meta-regression analysis of the same subset shows no significant effect, but this allows no conclusions. First of all meta-regression analysis should not be performed on good quality trials. Secondly, meta-regression on a relatively small number is imprecise. One could say that meta-regression of good quality trials should extrapolate to zero result, but we can neither conclude that this result confirms the efficacy of homeopathy. As stated, the number of good quality trials is too low to give reliable results, but there is also much heterogeneity involved. The 21 good quality trials consider 16 different

https://doi.org/10.51910/ijhdr.v8i28.352
indications. One indication, muscle soreness, is over-represented with four trials. Homeopathy is probably not effective for muscle soreness after marathon [13]. This is not surprising because marathon runners are healthy people. The four comparable conventional trials for this indication showed also no result, but none of these trials had good quality.

Figure 3: Comparison of 110 homeopathy trials (above) and 110 matched conventional trials (below) (source Shang, Egger et al; Lancet 2005; 366:726-732, with permission from Elsevier).

Meta-regression analysis could also be performed on all available trials. Comparative meta-regression analysis is possible if quality is comparable. The Lancet paper showed a suggestive figure where meta-regression lines of the homeopathy trials and conventional trials were compared. The regression line for homeopathy inclined more, but not significantly, towards zero effect. According to the authors this supported the hypothesis that homeopathy is a placebo effect, see Figure 3. But why is there a difference in asymmetry and what does it mean?

The comparison of regression lines between homeopathy and conventional medicine (lower part of Figure 3) is meaningless because quality (especially in smaller trials) is incomparable, but there is more.

There are several objections against the conclusion that this confirms the placebo hypothesis. First there is selection bias. Four trials in the left-upper quadrant of the homeopathy trials (upper part of Figure 3) were missing, because "no matching trials could be found". For one of them this is unexpected because it was a trial on rheumatoid arthritis.
The effects sizes of the two methods are incomparable because of the difference in publication bias and quality. Homeopathy has 16 (15%) unpublished trial and conventional medicine none, so the effect of conventional medicine is over-estimated. Low quality also leads to over-estimation of effects: inadequate concealment of treatment allocation is the most important quality factor, associated with 41% (95%CI: 27 to 52%) exaggeration of effect [14]. The difference of 12 good quality studies in favour of homeopathy, together with the difference in publication bias in 16 trials, leads to over-estimation of effect (compared with homeopathy) of one quarter (28) of all trials.

The last problem with this comparison is safety. Homeopathy is highly valued because of its safety and Shang et al disregarded this issue. The conventional study on weight loss showed a considerable positive effect of Dexfenfluramine [15], but Dexfenfluramine for weight loss was withdrawn by the American Food and Drug Administration in 1997 because of serious cardiac complications [16]. Two other larger studies, Deladumone (androgen-estrogen) in breastfeeding and Piroxicam for soft tissue injury suffered from the same problem [17-18]. These two treatments were also withdrawn because of adverse effects [19-20].

**Discussion**

The negative conclusion of the 2005 Lancet analysis of homeopathy was not rectified and this has been obscured for a long time because essential data were missing. Valid conclusions about effect of homeopathy compared to conventional medicine are impossible because of incomparability. The only valid conclusion is that quality of homeopathy trials, and especially of smaller trials, was better in homeopathy.

Since 1991 the discussion about proof for homeopathy is gradually shifting. The original question was: is homeopathy as a method a placebo effect? The outcome of about hundred trials on more than 30 indications is not compatible with the placebo hypothesis. Originally RCT evidence was demanded to overrule the implausibility of the method, but now things appear to be more complicated because of implausibility. Demands are becoming more specific: homeopathy for specific indications, homeopathy trials with good quality. In the end we have to deliver proof with good quality trials for each medicine for each indication. The effort to achieve this grows exponentially and with the present funding and infrastructure this is absolutely impossible. This discussion is no longer about placebo effect and statistical uncertainty, but about paradigms and trustworthiness.

The demand to investigate homeopathy as if there is no difference with conventional medicines is called 'hermeneutic injustice' by the philosopher Miranda Fricker [21]. Homeopathy cannot and does not work like conventional medicines. The effect for one medicine is not specific for one indication and there is no specific homeopathic medicine for one indication. Not recognising this difference is hermeneutic injustice. Compare this with Pre-menstrual Syndrome (PMS): if this diagnosis is not recognised, no treatment can be developed.

The 2005 Lancet paper goes beyond criteria for recognising conventional methods: there is a statistically significant effect even if only good quality trials are considered and quality is better than in matched conventional trials. But there are several sorts of bias known in conventional medicine that still could be considered, like incomplete reporting [22], fraud [23], and ghost authorship [24].

In 1846 the way a homeopathic medicine is manufactured was declared “an outrage to human reason” [25]. Since then Introduction of most analyses of homeopathy start with statements like "Many people assume that any effects of homeopathy must be non-specific placebo effects", without discussing the validity of this assumption in the light of new scientific developments like quantum mechanics, electromagnetism, etceteras [26]. Many of these developments would have been considered an outrage to human reason in 1846. The authors of epidemiological analyses are no experts in the field of possible mechanisms of action nor familiar with literature on the subject.
The paradigm is that homeopathy cannot work and beliefs cannot be overthrown by a single trial, however perfect. Bayesian philosophy is proposed to explain the persisting rejection of homeopathy [27]. But Bayesian philosophy is about updating beliefs after each new observation, and after five to ten trials even highly unlikely hypotheses become likely to probable [28]. If belief were growing with an increasing number of trials resistance should subside and funding and infrastructure for research should grow. Bayesian philosophy can also be applied in another way for differential diagnosis: does proof for homeopathy prove an intrinsic effect of homeopathy or bad science by homeopathic researchers? The process starts with a prior assumption: is the way a homeopathic medicine is manufactured really an outrage to human reason? In other words: are other mechanisms than chemical interactions impossible? Is, say, a resonance phenomenon triggered by electromagnetic signals impossible?[29] After establishing the prior probability of a mechanism other than conventional medicines updating of this probability can start using all available evidence. This evidence should also comprise expert opinion of patients and homeopathic doctors, fundamental research etceteras.

The key problem “is homeopathy really an outrage to human reason?” should definitively be resolved and the scientific community should be really interested to know the truth or else further research to prove homeopathy might be useless. Proof will be regarded as proof for bad science due to limitations of scientific research and all other kinds of proof that are permitted for conventional medicine will not be permitted for homeopathy [30].

Should the scientific community really be interested to know the truth about homeopathy? Epidemiological evaluation shows that antibiotics do not reduce the risk of serious complications in Upper Respiratory Tract Infections (URTI) [31-32]. One of the conclusions of the Lancet 2005 analysis was "The eight trials of homeopathic remedies in acute infections of the upper respiratory tract that were included in our sample, the pooled effect indicated a substantial beneficial effect (odds ratio 0.36 [95% CI 0.26–0.50]) and there was neither convincing evidence of funnel-plot asymmetry nor evidence that the effect differed between the trial classified as of higher reported quality and the remaining trials". A typical Cochrane review includes six trials [33]. Homeopathy could therefore be a realistic solution for a serious problem in the most frequently occurring acute disease in general practice.

Will the scientific community be interested? Kuhn states in The Structure of Scientific Revolutions: “… scientists fail to reject paradigms when faced with anomalies or counterinstances” (p. 78) [34]. An example: In a meta-analysis of statin treatment and the occurrence of haemorrhagic stroke Vergouwen et al found an effect of OR=0.88 (95% CI: 0.78-0.99). In the case of statins the 95% confidence interval is below 1.0 because the pooled sample size is large, n=8832. The result in Shang’s conclusive subset was OR=0.88 (95% CI: 0.66-1.18). The effect-size is the same as in the statins case, but the 95% Confidence Interval comprises unity because the sample size is much smaller, n=1923. There was also more heterogeneity in homeopathy because the conclusive subset was about eight trials on eight different indications, including ‘muscle soreness’, but excluding URTI. If we increase the sample size in homeopathy by taking all good quality trials (according to Shang) into account, we find a significant result OR=0.76 (95% CI: 0.59-0.99), n=2651. The significant result is already achieved if we comprise 14 good quality trials with sample size above median. Nevertheless, Shang drew a false negative conclusion based on insufficient sample size. Based on this false negative conclusion Shang stated that the result of homeopathy in URTI could not be trusted.

In this discussion ‘not-significant’ is confused with ‘not-effective’ and the therapy is proclaimed not-effective while there is still insufficient research and therefore insufficient sample size. Another example is reducing the sample size by creating a subset. Subsets with insufficient sample size can be created by selecting single indications [35]. If homeopathy is effective for an indication, like URTI, you can chose a subset of a subset. Altunc et al stated that there is no effect of homeopathy in URTI by selecting the subgroup ‘Childhood and Adolescence’ [36].
The homeopathic community should develop a strategy how to deal with science. If science is used to prove homeopathy we must realise that proof by RCT might never be conclusive because of shortcomings of RCT and meta-analysis. For new RCTs we’d better make strategic choices about indications to be evaluated. The main problem will remain the unbalanced and probably outdated estimation of the prior probability of the mechanism of action.

We are neglecting science as a means to improve homeopathy. Frei showed that results of RCT can be improved by improving the homeopathic method [37]. A large number of entries in the homeopathic repertories might be misleading and this can be improved by scientific research [38].

Conclusion

The results of homeopathic RCTs not different from the results of comparable conventional RCTs, with better quality of homeopathic trials. Discussions about proof for homeopathy are in fact discussions about science. The prior assumption that homeopathy cannot work pervades all aspects of this discussion and is not properly evaluated in the introduction of most analyses.

References

[1] Kleijnen J, Knipschild P, ter Riet G. Clinical trials of homeopathy British Medical Journal 1991;302:316-323
[2] Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges L, Jonas W. Are the clinical effects of homeopathy placebo effects? - A meta-analysis of Placebo-controlled trials. Lancet 1997;350:834-843
[3] Linde K, Scholz M, Ramirez G, Clausius N, Melchart D, Jonas Impact of study quality on outcome in placebo-controlled trials of homeopathy. J Clin Epidemiol. 1999; 52: 631–636
[4] Vandenbroucke JP, Crean JM de. Alternative medicine: a “mirror image” for scientific reasoning in conventional medicine. Ann Intern Med. 2001; 135: 507-513
[5] Ezzo J, Bausell B, Moerman DE, Berman B, Hadhazy V. Reviewing the reviews. How strong is the evidence? How clear are the conclusions? Int J Technol Assess Health Care. 2001 Fall; 17(4): 457-466
[6] Shang A, Huwiler-Müntener K, Nartey L, Jüni P, Dörig S, Sterne JAC, Egger M. Are the clinical effects of homeopathic placebo effects? Comparative study of placebo-controlled trials of homeopathy and allopathy. Lancet. 2005; 366: 726-732
[7] Lüdtke R, Rutten AL. The conclusions on the effectiveness of homeopathy highly depend on the set of analyzed trials. J Clin Epidemiol. 2008; 61: 1197-1204. [Doi: 10.1016/j.jclinepi.2008.06.015]
[8] Rutten ALB, Stolper CF. The 2005 meta-analysis of homeopathy: the importance of post-publication data. Homeopathy. 2008; 97: 169-177
[9] Sterne JAC, Egger M, Davey Smith G. Investigating and dealing with publication and other biases in meta-analysis. BMJ 2001;323:101-105
[10] Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997; 315: 629-634
[11] Vandenbroucke JP. Letter to editor: Experts’ views are still needed. BMJ. 1998; 316: 469-470
[12] Rutten ALB, Stolper CF. Diagnostic test evaluation by patient-outcome study in homeopathy: balancing of feasibility and validity. J Eval Clin Practice. 2009 In Press
[13] Ernst E, Barnes J. Are homoeopathic remedies effective for delayed-onset muscle soreness? - A systematic review of Placebo-controlled trials. Perfusion. 1998; 11(1): 4-8.
[14] Schulz KF, Chalmers I, Hayes RJ, Altman D. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995; 273: 408-412
[15] Enzi G, Crepaldi G, Inelmen EM, Bruni R, Baggio B. Efficacy and safety of dexfenfluramine in obese patients: a multicenter study. Clin Neuropharmacology. 1995; 12: S173-178
A revisão em 2005 de *The Lancet* comprovou a qualidade superior dos ensaios homeopáticos. O que vem a seguir?

**RESUMO**

Diversas revisões passadas evidenciaram efeitos positivos para a homeopatia, mas em 2005, *The Lancet* publicou uma revisão afirmando que a homeopatia é efeito placebo. Essa revisão foi criticada por não revelar informação essencial. A reconstrução dos dados pós-publicação se opõe a essa conclusão negativa. De fato, a única conclusão retificada pela metodologia da revisão de 2005 é que a qualidade dos ensaios homeopáticos, especialmente dos mais pequenos, é melhor que a dos ensaios convencionais. A comparação do efeito de 110 ensaios homeopáticos com 110 ensaios compatíveis convencionais sofreu de vício de seleção, de publicação diferente, qualidade diferente e segurança diferente. Contudo, não há diferença significativa nos efeitos entre ambos os métodos. As discussões acerca das provas a favor da homeopatia são, na verdade, discussões sobre a ciência. O pressuposto de que a homeopatia não pode funcionar permeia todos os aspectos desta discussão e não é avaliado apropriadamente na introdução da maioria das análises.

**Palavras chave:** Homeopatia; Revisão; *The Lancet* 2005.

---

La revisión de 2005 de *The Lancet* comprobó la calidad superior de los ensayos homeopáticos. ¿O que sigue ahora?

**RESUMEN**

Varias revisiones anteriores demostraron los efectos positivos de la homeopatía, pero la publicada en 2005 por *The Lancet* afirma que homeopatía es efecto placebo. Esta última revisión fue criticada por no revelar información esencial. La reconstrucción de los datos pos-publicación contradice aquella conclusión. La única conclusión rectificada por la metodología utilizada en la revisión de 2005 es que la calidad de los ensayos homeopáticos, especialmente de los más pequeños, es mayor que la de los ensayos convencionales. La comparación del efecto de 110 ensayos homeopáticos y 110 convencionales apareados sufrió vicio de selección, vicio de publicación diferente, calidad diferente y seguridad diferente. Sin embargo, no hay diferencia significativa entre los efectos de ambos métodos. Las discusiones acerca de las pruebas a favor de la homeopatía son, en verdad, discusiones sobre ciencia. La presuposición de que la homeopatía no puede funcionar permea todos los aspectos de esta discusión y no es evaluada apropiadamente en la introducción de la mayoría de tales análisis.

**Palabras llave:** Homeopatía; Revisión; *The Lancet* 2005.