Asymmetry in *The Lancet* meta-analysis

We have read with interest the paper by Shang *et al.*[1] We appreciate the efforts to undertake such an enormous work. In the following we want to express some comments from a scientific point of view.

The exciting message is that homeopathy works, as described in the Results section: in both groups a beneficial effect was shown when all 110 studies were taken into account (Figure 2). The authors also acknowledge less pronounced heterogeneity for homeopathy trials. In addition, a higher quality of included studies was found in the homeopathy group (19 *vs* 8%, Table 2).

While agreeing with the authors with respect to these sensational results, we have major concerns about their conclusions.

Firstly, even with careful selection, it remains problematic to compare studies of a pool of 165 for homeopathy *vs* >200,000 for conventional medicine. This factor of >1000 already contains asymmetry. Furthermore, it appears that there is discrimination when publications in English (94/110, 85% in the conventional medicine group *vs* 58/110, 53% in the homeopathy group) are rated higher quality (Table 2).

Neither the Summary nor the Introduction clearly specify the aim of the study. Furthermore, the design of the study differs substantially from the final analysis and therefore the prolonged description of how the papers and databases were selected is misleading: instead of analysing all 110 studies retrieved by their defined inclusion and exclusion criteria, the authors reduce the number of investigated studies to 'larger trials of higher quality'. By using these sub-samples, the results seem to differ between conventional medicine and homeopathy.

The meta-analysis does not compare studies of homeopathy *vs* studies of conventional medicine, but specific effects of these two methods in separate analyses. Therefore, a direct comparison must not be made from this study.

However, there remains great uncertainty about the selection of the eight homeopathy and the six conventional medicine studies: the cut-off point seems to be arbitrarily chosen: if one looks at Figure 2, the data look very much the same for both groups. This holds true even if various levels of SE are considered. Therefore, the selection of larger trials of higher quality is a post-festum hypothesis but not a pre-set criterion. The question remains: was the restriction to larger trials of higher quality part of the original protocol or was this a data-driven decision? Since we cannot find this proposed reduction in the abstract, we doubt that it was included *a priori*.

However, even if one assumes that this was a predefined selection, there are still some problems with the authors' interpretation: for larger trials of higher reported methodological quality, the odds ratio was 0.88 (CI 95%: 0.65–1.19) based on eight trials of homeopathy: although this finding does not prove an effect of the study design on the 5% level, neither does it disprove the hypothesis that the results might have been achieved by homeopathy. For conventional medicine, the odds ratio was 0.58 (CI 95% 0.39–0.85), which indicates that the results may not be explained by mere chance with a 5% uncertainty. There may be 'evidence' or 'no evidence' but not 'weak or strong evidence'.

Although the authors acknowledge that 'to prove a negative is impossible' the authors clearly favour the view that there is evidence that homoeopathy exhibits no effect beyond the placebo-effect. However, this conclusion was drawn after a substantial modification of the original protocol which considerably weakens its validity from the methodological point of view. After acquiring the trials by their original inclusion- and exclusion criteria they introduced a further criterion, 'larger trials of higher reported methodological quality'. Thus, eight trials (= 46% of the larger trials) in the homoeopathy group were left and only six (32%) in conventional medicine group (an odds ratio of 0.75 in favour of homoeopathy).

But the decisive point is that it is unlikely that these six trials are still matched to the eight samples of homoeopathy (although each of the 110 in the original was matched). Consequently, one cannot conclude that these trials are still comparable. Thus, any comparisons of results between them are unjustified.

The rationale for this major alteration of the study protocol was the assumption, that these larger, higher quality trials are not biased, but no evidence or data-based justification is given. Neither the actual data (odds ratio, matching parameters...) nor a funnel plot (to indicate that there is no bias) of the final 14 trials are supplied although these parameters constitute the ground of their conclusion.

The other 206 trials (94% of the originally selected according to the protocol) were discarded because of possible publication biases as visualized by the funnel plots. However, the use of funnel plots is also questionable. Funnel plots are thought to detect publication bias, and heterogeneity to detect fundamental differences between studies. New evidence suggests that both of these common beliefs are badly flawed.

Using 198 published meta-analyses, Tang and Liu demonstrate that the shape of a funnel plot is

largely determined by the arbitrary choice of the method to construct the plot.[2] When a different definition of precision and/or effect measure was used, the conclusion about the shape of the plot was altered in 37 (86%) of the 43 meta-analyses with an asymmetrical plot suggesting selection bias. In the absence of a consensus on how the plot should be constructed, asymmetrical funnel plots should be interpreted cautiously. These findings also suggest that the discrepancies between large trials and corresponding meta-analyses and heterogeneity in meta-analyses may also be determined by how they are evaluated.

Researchers tend to read asymmetric funnel plots as evidence of publication bias, even though meta-analyses without publication bias frequently have asymmetric plots and meta-analysis with publication bias frequently have symmetric plots, simply due to chance. Use of funnel plots is even more unreliable when there is heterogeneity.[3]

Apart from the questionable selection of the samples there is a further aspect of randomness which further weakens their conclusion: the odds ratio of the eight trials of homoeopathy was 0.88 (CI 0.65–1.19), which might be significant around the 7–8% level. Actually, the reader might be interested to know at which exact level homeopathy would have become significant. Thus, there is no support of their conclusion any more when you shift the level of significance by mere, say 2–3%. In addition, with such controversial hypotheses the scientific community would tend to use a level of significance of 1% in which case the odds ratio of the conventional studies would not be significant either.

From a statistical point of view, the power of the test, considering the small sample sizes, should have been stated, especially in the case of a non-significant result. Above all, the choice of which trials are to be evaluated is crucial. By choosing a different sample of eight trials (eg the eight trials in 'acute infections of the upper respiratory tract', as mentioned in the Discussion section) a radically different conclusion would have had to be drawn (namely a substantial beneficial effect of homeopathy—as the authors state). The authors may not be aware that larger trials are usually not 'classical' homeopathic interventions, because the main principle of homeopathy, individualization are difficult to apply in large trials. In this respect, the whole study lacks sound understanding of what homeopathy really is.

At the Medical University of Vienna, we tend to investigate controversial issues on an academic basis. We support the dialogue between conventional medicine and homeopathy. We hope that Lancet will be open-minded for non-conventional medicine in the future as it was in 1994 and 1997.[4,5]

References

1 Shang A, Huwiler-Muntener K, Nartey L, *et al.* Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. *Lancet* 2005; **366**: 726–732.
2 Tang J, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *J Clin Epidemiol* 2000; **53**: 477–484.
3 Terrin NC, Schmid CH, Lau J. Publication bias, chance, and heterogeneity: how researchers interpret the funnel plot. *Ninth International Cochrane Colloquium*, Lyon, France, 9–13 October 2001. *Cochrane* 2001; **1**: op006.
4 Reilly DT, Taylor MA, Beattie NGM, *et al.* Is evidence for homoeopathy reproducible? *Lancet* 1994; **344**: 1601–1606.
5 Linde K, Clausius N, Ramirez G, *et al.* Are the clinical effects of homoeopathy placebo effects? A meta-analysis of placebo-controlled trials. *Lancet* 1997; **350**: 834–843.

Michael Frass
Department of Internal Medicine I Medical University Vienna and Ludwig Boltzmann Institute for Homoeopathy, Vienna, Austria
E-mail: michael.frass@meduniwien.ac.at

E Schuster
Core Unit of Medical Statistics and Informatics, Medical University Vienna, Vienna, Austria

Ilse Muchitsch
Interdisciplinary Homoeopathic Research Group, Austrian Pharmacists Association, Austria

Jeff Duncan and Walter Geir
Doctors Association for Classical Homoeopathy

Gloria Kozel, Christa Kastinger-Mayr and Anton E Felleitner
Austrian Society for Homoeopathic Medicine, Austria

Christian Reiter
Umbrella Organisation of Austrian Doctors for Holistic Medicine, Austria

Christian Endler
Interuniversity College for Health and Development, Graz, Austria

Menachem Oberbaum
The Center for Integrated Complementary Medicine Shaare Zedek Medical Center, Jerusalem