Reviewer's report

Title: Method for appraising model validity of randomised controlled trials of homeopathic treatment: multi-rater concordance study

Version: 1 Date: 2 January 2012

Reviewer: Harald Walach

Reviewer's report:

This paper is well written and important. Methodologically it is sound and represents a well conducted exercise of validating a system to rate model validity of trials that report on homeopathic studies. It is a much needed piece of work and I would support publication by all means.

It is also well presented and gives little reason for critique. However there is one, which I would like to point out and would urge the authors to take up as a paragraph in the Discussion section. I would classify this as a minor compulsive change which needs to be made before publication, but which editors can check for completion; I need not (and do not want to) see the paper for this again, as it is simple and straightforward.

The authors say that the results did not play any role in their considerations. I find that this is probably as they might want reality to be, but it is hardly credible that this is the case. I find the case in test is reference 26, Lewith at al’s replication of the previously successful Reilly et al. study published in Lancet. I can recall the homoeopathy community’s jubilation on that publication. No one that I know of would have taken up their pen and then remarked that, sorry, but this is not homeopathic practice. Only after the negative trial of Lewith and colleagues was published was this argument raised. While the argument itself, strictly speaking, is true in that the intervention tested is isopathy, and isopathy is one specific form of homeopathy which is sometimes usef for allergies, it is also irrelevant in this case. For to ask, whether isopathy is a good form of homeopathy is to ask homeopaths of one lineage, whether homeopathy of another lineage is a good form of homeopathy. The answer one gets will depend on the lineage of the homeopathy, and, more importantly, on the outcome of the study itself.

While this argument does not invalidate the work of the group, it does invalidate some of the actual ratings obtained, since it is very hard to believe that the same ratings had been obtained, had the study been blind to results. Since it is impossible to blind raters of a field so small as homeopathy research to outcomes, it is for practical reasons impossible to do the study otherwise than with results already obtained and hence unblinded to outcomes. But I feel, it would be necessary and fair to acknowledge this shortcoming in a paragraph in the Discussion section. Ideally, one would have wanted to do such a validation exercise blind to outcome such that results do not taint the picture. This would have only be possible prospectively on protocols of studies whose results have
not been obtained yet. As this is unrealistic, the authors chose the second best way, but that creates a weakness of the procedure, not necessarily of the outcome. The domains are clearly relevant, and the agreement rating is good enough. But it might have been even better, had the ratings been done blind to outcome.