Author's response to reviews

Title: Evidence for placebo effects on physical but not on biochemical outcome parameters. A review of clinical trials

Authors:

Karin I Meissner (karin.meissner@med.uni-muenchen.de)
Hans Distel (hans.distel@med.uni-muenchen.de)
Ulla Mitzdorf (mitzdorf@med.uni-muenchen.de)

Version: 2 Date: 17 February 2007

Author's response to reviews: see over
Dear Editor,

Herewith we send you the revised manuscript to BMC Medicine, entitled “Evidence for placebo effects on physical but not on biochemical outcome parameters. A review of clinical trials”.

We would like to thank the reviewers for their favorable comments and constructive suggestions. We have carefully addressed each of the reviewers' concerns and provide a point-by-point response to the comments below.

Best regards,

Karim Meissner, Hans Distel and Ulla Mitzdorf

Response to Reviewer 1 (Klaus Linde):

General
This is a very interesting paper on an important topic. The methodological approach is straightforward and adequate but, necessarily, not without problems. Selecting placebo-controlled trials of "stable" conditions for estimating the size of placebo effects (part 1) is difficult and cannot be made fully transparent within the space limits of a manuscript. Subgroup analysis of a published meta-analysis (part 2) also does not have perfect conclusiveness. These shortcomings are adequately discussed and difficult to avoid. I have one fundamental point, but as this depends on my personal view of the "placebo problem" I have put this under discretionary revisions.

---------------------------------
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
None
---------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

- What is missing in the discussion are more detailed recommendations for future research. This paper is a good step in the direction of looking for reasons for the heterogeneity of placebo effects (see below).

We have now added a discussion paragraph, in which we refer to the heterogeneity of placebo effects and to possible implications for future research.

- page 10: para 1 and para 2 - number of trials 26 vs. 29
We have corrected the number of trials to 26.

Discretionary Revisions (which the author can choose to ignore)
Personal, fundamental point: In my view the main merit of this paper is that it provides an empirical investigation of a plausible differentiating hypothesis (that placebo interventions are not the same for all outcomes) in an area where very heterogeneous things are often lumped together. What I do not like, however, is that the authors still seem to consider "placebo" as a homogeneous intervention (at least, this is suggested by terms such as "THE placebo phenomenon" or "placebo treatment"). In my view it is much more likely that "placebo effects" are a class of effects with a number of different mechanisms. Depending on the circumstances (patients, condition, interventions, outcome etc.) different mechanisms might be relevant. The concept of "THE placebo effect" might simply be nonsense. I would suggest to at least use terms such as "placebo phenomena" and "placebo interventions" which imply that there could be heterogeneity.

We agree with the reviewer’s opinion that the term “placebo effect” may wrongly suggest a unique mechanism to be involved. In addition, we agree that placebo treatment cannot be considered a homogenous intervention, because, depending on the circumstances, different placebo-generating factors may be involved. We therefore have replaced the terms “placebo effect”, “placebo treatment” and “placebo intervention” by their plurals throughout the manuscript wherever possible, as suggested. In addition, we have replaced the term “placebo phenomenon” in the abstract by “placebo effects in clinical trials”. We also refer now to this issue of heterogeneity in the discussion section.

Minor point: page 11 - "In comparison to the pharmacological medication ... by one third -a remarkable efficacy”. Frankly, I was surprised that it was ONLY one third, and that made me wonder about how reliable the primary studies are (unblinding etc).

Compared with the placebo efficacy rate reported in other fields (e.g. about 75% in depression) a rate of one third may appear quite low, and patient unblinding, indeed, may have diminished the effect. However, the aim of the present study was to investigate whether peripheral outcome parameters respond to placebo treatment, and in this sense an average rate of one third for placebo effects on physiological outcome parameters in relation to pharmacological effects appears quite remarkable to us, since it clearly indicates the existence of peripheral placebo effects.
Response to Reviewer 2 (A. Jon Stoessl):

General
The authors have performed 2 meta-analyses of studies in which the placebo effect was assessed. They conclude that in peripheral organs, the placebo effect is more likely to be detected when a “physical” outcome measure is used, as opposed to biochemical measures, which seem to be resistant to placebo effects (or possibly even negatively impacted). For the first dataset, the authors identified 34 studies in 17 different conditions from 1689 double blind randomized trials investigating chronic stable clinical conditions in which at least one peripheral outcome parameter was reported. Of these, another 5 trials were excluded because of insufficient data. They then determined the significance of changes from baseline within the placebo group of each trial. In this analysis, significant placebo effects were found for a variety of physical parameters, such as blood pressure, pulmonary flow rates, urinary flow rates or residual volume, but not for biochemical measures such as hormone levels, enzymes, inflammatory proteins, HbA1c or cholesterol. Interestingly, the ESR was subject to placebo effects in a study of rheumatoid arthritis. In studies with ‘physical’ outcome measures, a significant placebo effect was seen in 8/16 trials, in contrast to only 1/18 trials with biochemical outcome measures. The effect size for studies with physical outcome measures was approximately 1/3 of the value seen in the best active treatment arm.

In order to confirm their findings with an independent dataset, the authors then examined the studies included by Hrobjartsson & Gotzsche in their meta-analysis comparing the effects of placebo against no treatment. Whereas Hrobjartsson & Gotzsche had concluded that there was no evidence for a placebo effect, Meissner and her colleagues, selecting 26 trials with peripherally measured outcomes, found a significant placebo effect when the analysis was restricted to physical outcomes, but if anything a negative effect of placebo when biochemical outcome measures were examined.

This is an interesting report that lends further support to the importance of placebo effects, even in conditions other than neuropsychiatric disease, where these are widely recognized. It is of particular interest that the widely publicized analysis of Hrobjartsson & Gotzsche does in fact reveal significant effects of placebo when subjected to the subgroup analysis applied by Meissner and her colleagues. The authors suggest that their ability to detect ‘physical’ but not biochemical placebo effects reflects visceral reward learning, dependent upon contingent feedback to cortical areas involved in autonomic control. It may seem somewhat arbitrary to differentiate between physical and biochemical outcomes and in the case of neuropsychiatric disease, it is clear that at least some of the effects of placebos are mediated by neurochemical changes – e.g. dopamine release in the case of Parkinson’s and pain, opioid release in the case of pain. Thus the crucial difference is that the majority of the biochemical measures were taken in conditions where the time course may be protracted and where there
may not be any associated feedback – thus, patients do not know what to expect in response to therapy and it is this expectation that is critical to the placebo effect.

Thank you for these interesting remarks. With respect to biochemical changes in neural diseases, we would like to argue that neurotransmitters, e.g. acetylcholine and norepinephrin, will most probably also be involved in the mediation of autonomic changes leading to placebo effects on peripherally-measured physical outcomes. However, in our view, neurotransmitters should be regarded a special class of biochemical parameters not necessarily comparable with the kind of peripherally measured biochemical outcome parameters we addressed in our study.

---

**Major Compulsory Revisions** (that the author must respond to before a decision on publication can be reached)

No major compulsory revisions required.

---

**Minor Essential Revisions** (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. There is an error in Table 2 (supplemental material on the first dataset): Ref. 19 is a study on hypertension, not diabetes. The same mistake occurs in Fig. 1. Also in Table 2, Ref. 36 should refer to diabetes, not hypertension. Many of the entries on Figure 2 do not match up with the references or entries in Table 3. Similarly Ref. 63 (Table 3) refers to hyperprolactinemia, not diabetes, whereas the converse is true for Ref. 67.

Thank you for identifying these errors. We have checked all tables and references and have corrected all mis-entries.

2. Pg. 10, 2nd paragraph, first line should read “The meta-analysis of all 26 (not 29) trials…” .

*We have corrected the number of trials to 26.*

---

**Discretionary Revisions** (which the author can choose to ignore)

1. Although not explicitly stated in the paper, the authors presumably chose to exclude neurological and psychiatric disorders precisely because they are recognized to be highly susceptible to placebo effects, even when objective outcome measures (both physical, such as motor scores or electrophysiological recordings and biochemical, such as measures of neurotransmitter release) are
used, and because one might expect a more tenuous link between expectation and ‘peripheral’ outcomes.

Our main reason for focusing on peripheral disease processes was the rather scarce evidence for placebo effects in this field. We mention this now in the introduction.

2. While the authors have found one explanation for why Hrobjartsson & Gotzsche failed to validate a placebo effect in their meta-analysis, there are numerous other reasons that the conclusions of the Danish investigators should be viewed with caution. First, they included in their analysis studies performed in subjects with questionable ability to generate expectation. Secondly, the degree of expectation was low (placebo was compared to either no treatment alone, or to at most a single active treatment arm out of 3 arms). Thirdly, binary outcome measures (improved or not) were used, thus a graded degree of improvement would not be detected.

We fully agree that there are also other possible explanations for the lack of a placebo effect in the meta-analysis of Hrobjartsson & Gotzsche, among them studies with low levels of expectation, and the use of insensitive outcome measures in the analysis of binary outcomes. We therefore have changed the expression “A possible reason…” to “One possible reason…”.

3. The analysis of changes compared to baseline is fraught with difficulty in the absence of a no treatment group, as apparent benefit could occur in the absence of a true placebo effect, reflecting regression towards the mean. The authors are aware of this and attempted to account for this confound by analyzing the second dataset, but cannot confidently exclude this possibility in the first dataset. A more robust approach to the analysis might be to apply analysis of covariance with baseline values as the covariate, rather than to simply rely on change from baseline.

We agree with the reviewer’s opinion that analysis of covariance would have been the better choice to analyze the placebo effects in the trials of our first dataset. However, to our best knowledge, analyses of covariance can only be computed when either individual patient data, or at least the results of analyses of covariance for individual trials are available. Since these data were not provided, we could only use standard meta-analytic procedures to estimate the average placebo effect. Nonetheless, we refer now to the problem of regression towards the mean in more detail in the discussion section.

4. The authors’ choice to include acute or post-surgical conditions in the second dataset may make the interpretation of their results difficult, as the magnitude (and mechanism) of the placebo effect may differ between acute and chronic conditions. In this dataset, they chose to include one trial in which subjects had
high normal blood pressure. Although this condition may be highly sensitive to regression towards the mean, it would probably be associated with a negligible placebo effect, as subjects would have very little motivation to ‘improve’. Another potential bias is derived from the authors' decision to restrict their analysis to “the parameter or measurement with the largest effect size during the treatment period” in situations where numerous parameters or time points had been measured.

**We gratefully picked up the notion of a possible difference between placebo effects in acute/ postsurgical and chronic conditions and added a sensitivity analysis of the second dataset with acute/ postsurgical conditions excluded. Interestingly, this did not significantly alter the results.**

We agree with the reviewer’s notion that our inclusion of a trial investigating subjects with highly normal, or ‘borderline’ blood pressure might be debatable, since these subjects may have little motivation to improve. However, high level of expectation was not a selection criterion for the trials of our study, and sensitivity analysis with this trial excluded did not alter the results substantially.

**Our choice to analyze parameter changes and time points displaying the always largest placebo effects may have produced some bias. However, our primary study aim – to identify parameters with differential placebo responsiveness – most probably was not affected by this procedure.**

**5. The authors suggest that negative effects of placebo in the second dataset might be related to dietary factors. Another possibility is that participation in a study where patients feel they have a low likelihood of deriving benefit, or even more so, should they become unblinded and find that they have been receiving placebo, might lead to negative placebo effects. Whether or not this is more likely to occur with biochemical outcomes is not clear. The significance of the negative placebo effects on biochemical outcomes is in any event questionable. In the first dataset, there is a very large study (N>900) of cholesterol in which there was no effect of placebo.**

**We fully agree with the reviewer’s opinion that the negative effect should be viewed with caution, and we also mentioned this in the discussion. The suggestion that low expectation, or even unblinding of placebo recipients may lead to negative placebo effects is plausible; we added this point to the discussion, thereby supposing that the moderating variable may also be compliance with the dietary regimen.**
Response to Reviewer 3 (Ted J. Kaptchuk):

General
This reviewer enjoyed reading this thoughtful, well-performed and well-written study. The question addressed, to what extent does placebo treatment effect physically defined measures of peripheral disease process as opposed to biochemical parameters is important and timely. This reviewer has several minor questions and comments.

1. In abstract, in the conclusion, the discussion of visceral reward learning should not be there and rather confined to the discussion. The points are important but this discussion is an interpretation of results and not a conclusion for abstract.

We have replaced the discussion of visceral reward learning by a remark on future investigations.

2. page 4. Medline does include “untreated control groups” but there is no regular or established way to search for this condition.

This suggestion is very helpful for future reviews. We slightly changed the respective sentence to make clear that the lack of untreated control groups was not due to performing a Medline search, but to our own research strategy.

3. page 5. The biggest concern I have is what is called “stable disease.” What exactly is meant seems to be judgment call and might be defined with some examples. See below.

We have inserted now in the methods section two examples for illustrating which conditions were judged to be ‘stable’. We would like to emphasize that the selection process was performed by using exclusion rather than inclusion criteria. Thus, a condition was judged to be ‘stable’ when we could not identify any other factor than placebo treatment that obviously influenced the course of disease in the placebo groups (see also Table 1 for a summary of exclusion criteria). Since we cannot fully exclude that regression toward the mean, or symptom variation may have influenced the course of otherwise stable diseases, we gave these limitations more emphasis in the discussion.

4. page 6. Besides unblinding, I think a second reason to only first phase of cross-over is that there may be a carry-over effect from the drug.

Thank you, we have added this second reason to the methods section.

5. page 7. Please check the sentence that concerns 56 trials and only 13. Is there an error here?
The sentence did not express well what we intended to say. For better comprehensiveness we have elaborated it: “First, we selected all 56 trials with observer-reported continuous outcomes. These included 13 trials with corresponding patient-reported and observer-reported continuous outcomes, from which only the patient-reported ones had been included in the main analysis by Hróbjartsson and Gøtzsche.”

6. Discussion. I’d make the caveats and limitations stronger. Again, what is stable disease? For example, fistula is considered stable. But some the improvement in physical could be observer bias. Also, I think many of these conditions are subjected to variability and regression. If the definition or caveats were clear, I think the category of “stable” would be ok.

We have now added a paragraph in the discussion, in which we refer to these caveats and limitations in detail.

With respect to the suggestion of possible improvement of physical parameters by observer bias, we would argue quite contrariwise: In classical placebo-controlled trials, an observer’s knowledge of the intervention when assessing treatment effects could lead to overestimating the effect in the active treatment arm but probably not in the placebo arm. Due to the general low expectation of scientists with respect to physiological placebo effects, observer bias may even have led to underestimate the placebo effects.

7. Good discussion of operant condition and summary of recent studies.

Thank you, we are delighted reading this!

**Major Compulsory Revisions** (that the author must respond to before a decision on publication can be reached)

**Minor Essential Revisions** (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

**Discretionary Revisions** (which the author can choose to ignore)