Replies to the reviewer comments on:

Title : Zinc lozenges may shorten the duration of colds: a systematic review
Authors : Harri Hemilä
MS ID : 6567020912717028
Journal : BMC Infectious Diseases
Article type: Research article

Submitted April 28, 2009 to BMC Infectious Diseases

Reviewer comments received by Hemilä at August 17, 2009

****

Hemilä's replies (HH) to Reviewer 3 comments Oct 21, 2009:

Referee 3:

http://www.biomedcentral.com/imedia/1956627467294205_comment.pdf

Reviewer's report
Title: Zinc lozenges may shorten the duration of colds: a systematic review
Version: 1 **Date:** 28 July 2009
Reviewer: Lehana Thabane

Reviewer's report:

1. Abstract: Background
a. The purpose of the study is stated as "to examine whether the total dose of zinc might explain part of the variation in the results". This seems to suggest that the primary aim of the review is to perform a subgroup analysis. This is rather uncommon as subgroup analysis is usually secondary.

HH: There is nothing inappropriate in formulating an objective that I am primarily interested in the dose-dependency, motivated by the earlier literature to which I refer in the Introduction.

b. The primary aim of the review should clearly stated using the PICOT (population, intervention, comparator, outcome, time frame) format.

HH: I do not understand this comment. Even the title states the intervention and outcome of primary interest. If I would restrict to adults, that could be added to the title, but it is inefficient use of space to add to title that the review covers both adults and children.

2. Abstract: Methods
a. Why was combining p-values chosen for analysis?

HH: There are two general goals in medical statistics:
1) to test the hypothesis that there is difference between two groups and
2) to estimate the size of the possible difference between the study groups.
Sometimes these two goals can be reached simultaneously e.g. when calculating a pooled RR value in which case the confidence interval gives answer to 1). Pooling the results, however, requires that the conditions of the studies are "close enough".

Testing hypothesis by e.g. the Fisher method is much less sensitive to the variations in original studies. For example, one study might report the duration of colds in days (+/- SD) and another might report the duration as the number of persons who have colds longer than 1 week. Such results cannot be combined to a pooled estimate of effect because the outcome definitions are so different. However, both studies test the effect of intervention on duration of colds and the resulting P-values can be combined by the Fisher method.
Thus, my reply to reviewer's question is: I am testing the hypothesis that there is difference between zinc and placebo groups in the selected studies.

b. State clearly if the analysis used fixed- or random-effects approach.

HH: The last sentence of my Method section stated clearly: "I combined … by using the RevMan program (version 5) using the inverse-variance fixed-effect option [34]."

3. Methods: Search of the trials
a. The search should be done in duplicate by two independent reviewers. This will allow readers to assess whether it could be reproducible.

HH: First, carrying out the search "in duplicate" is no proof that the search is sound or "reproducible." Two and more people can do poor job (I have examples). Second, one person can search the literature carefully.

b. Where the MESH equivalents of the search terms used? If not, why not?

HH: Ovid Medline gives 76404 records when I search for "zinc" as the free search term (.mp). If I search zinc as the MESH term, I get 40730 records. The latter 40730 records are all included among the former 76404 records. Thus, there are no records missing with the use of zinc as free search term.

c. Why were other common databases such as EMBASE not searched?

HH: I do not have free access to EMBASE, which was thus a pragmatic reason not to use EMBASE. However, there are lots of other data bases that I have not searched either. It is always possible to speculate that there might be some trial(s) published in such an obscure journal that it is not identified through MEDLINE.

On the other hand, I used also the CENTRAL, which covers the trial literature widely and, in addition to collecting data of controlled trials from several data bases, there has been lots of hand searching of relevant journals to make the CENTRAL a good resource for identifying controlled trials.

I also used SCOPUS. EMBASE and SCOPUS are both data bases of Elsevier. I assume they have substantial overlap, but I have not been able to find exact description of the overlap.

d. Were there any language restrictions?

HH: No language restriction was described in the Methods, which implies that there was no language restriction. This was added to the Methods as an explicit statement.

e. How was the triple counting of placebo group patients dealt with for the trial that had 3 groups?

HH: The placebo group of Turner's four-arm trial is counted three times in the "all trials" line of Table 2. It is counted one time in the line "Zn dose over 75 mg/day" and two times in line "Zn dose less than 75 mg/day". The same placebo group makes the three study arms non-independent in the "all trials" group.

However, given that all three comparisons have P=0.5, counting the Turner's placebo group three times is a conservative approach. I could combine all the zinc arms to a single pooled zinc group and compare that to the single placebo group. In this approach we would have only one P=0.5 remaining in the combination of P-values, which would lead to an even smaller combined P-value, meaning even stronger evidence that zinc differs from placebo.

In the low-dose subgroup, the combined P-value indicates lack of effect, and this lack of effect is not influenced by the Turner study. We can leave out one or both of the Turner arms, and the fining is the same.

In the high-dose subgroup, the Turner's placebo group is counted only once. Thus, the use of the single placebo group against the three zinc arms in the Turner trial has no effect on the conclusions.

4 .Methods: Calculation of the daily zinc dosage
a. One assumes that the calculation for zinc dosage was done to address the subgroup analysis of whether the total daily zinc could explain the heterogeneity. However, there is hypothesis and corresponding rationale stated for the subgroup analysis.

HH: I do not understand this comment.

5. Methods: Statistical methods
a. It is unclear why the pooling of p-values was used instead of pooling the effect estimates.
HH: This is explained above.

b. The data abstraction should be done in duplicate with agreement between reviewers measured using Kappa statistics.

HH: There is no objective basis for such a statement. I can myself check and re-check the numbers that I am extracting from the original papers. In some cases there are subjective aspects in the interpretation of study reports, but there are no subjective interpretations in the extractions of the numbers to Table 1.

c. There is no clearly stated research hypothesis/objective/question.

HH: The last sentence of the Introduction states: "The purpose of this systematic review is to examine the relationship between the total daily dose of zinc from the lozenges and the effect of the lozenges on patients who had natural common cold infections."

d. There is no assessment of study quality and potential impact of study quality on findings.

HH: At the beginning of the Results, I write: "All these trials were double-blind, although this feature was not used as the selection criterion. Weissman et al. [27] did not report the method of allocation, but all other trials were randomized." This is assessment of study quality and thus the reviewer's above comment is incorrect.

At the end of the Results section, I write "I did not carry out subgroup analysis by the methodological quality of the trials because all trials were double-blind trials. The only one not reported as randomized did not find any benefit [27]. Consequently, its exclusion would strengthen the evidence that zinc lozenges differ from placebo. In particular, all three trials which used zinc acetate in doses higher than 75 mg/day were methodologically rigorous randomized trials [20,21,24]." This is assessment of potential impact of study quality and thus the reviewer's comment above is incorrect.

e. The sensitivity of the assumption of using p=0.5 in the analysis needs to be

investigated.

HH: When there is insufficient data to calculate the accurate P-value and no P-value is reported, it is safe to be conservative and assume that zinc equals placebo, which means P=0.5. There is no justification to use, say, P=0.9 or P=0.1 in the studies that did not report suitable data for accurate calculation of P value.

6. Results
a. Somewhere (no page number), the author states that "I did not carry out subgroup analysis by the methodological quality of the trials because all trial were double-blinded trials". It is important to note that blinding is not the only thing that affects reporting quality of trials. Lack of reporting of allocation concealment and intention-to-treat analysis has also been associated with biases of effect estimates.

HH:
First, a meta-analysis of 276 randomised controlled trials concluded that "our analysis did not reveal any consistent associations between quality measure and the magnitude of the treatment effect in 4 clinical areas. In particular, double blinding and allocation concealment, 2 quality measures that are frequently used in meta-analyses, were not associated with treatment effect" (Balk et al. JAMA 2002;287:2973-82; Ref. 48 of the current manuscript.). Thus, there is empirical evidence against the importance of "allocation concealment" as a universally crucial quality measure.

Second, "allocation concealment" means that the participant and others directly involved in treatment do not know to which treatment group the participant had been allocated. This issue can be crucial in therapeutic trials in which the severity of a disease may affect allocation if it is not concealed. Even if the treatment cannot be blinded (e.g. in surgery), it is possible to carry out allocation as concealed, and therefore the concealment of allocation is a relevant issue independent of blinding of the intervention and outcome assessment. However, when the trial is "double-blind" that means that allocation must have been concealed. It is not possible to carry out the trial as "double-blind" if the allocation was not concealed. Therefore, double-blinding means that the trial must have been using concealment of allocation.

Finally, in the end of Results, I write "In the Petrus trial, only one participant was lost from follow-up [21]; in the first Prasad trial, two participants in the placebo group dropped out on day 2 [24], whereas there were no drop-outs in the second Prasad trial [20]." Thus, I considered the ITT question, and given the low number of participants who were lost from follow-up it is not reasonable to suggest that the differences in Table 3 are biased by drop-outs.

7. Other:
a. The paper needs serious editing for English grammar.

HH: The text had been checked by a professional linguist at the Language Center of our University.

b. It is important to include a flow-diagram of the search and selection of studies for the review as recommended by the QUOROM Statement.

HH: In some cases a flow diagram can be useful. However, a manuscript is a compromise between the goal of a rather short and easy to follow text (number of words and tables and figures) and the presentation of all relevant details. Essentially always there are lots of details that are relevant to the study, but they must be left out to keep the text readable. The flow-diagram would take one page of a printed paper, while the information provided to a reader is minimal compared with the space.

I added a flow chart as an Appendix.

c. Consider using correct plots to display the results.

HH: When there is a large number of trials that are not clinically or statistically heterogeneous, a forest plot is an economic way to show the distribution of results and their pooled effect. In Table 3 there are only three trials and presenting the estimates and confidence intervals gives the reader accurate information in contrast to a plot.

Quality of written English: Not suitable for publication unless extensively edited

HH: The text had been checked by a professional linguist at the Language Center of our University.