Previous reviewer comments:

Journal Name: Clinical Epidemiology

Title: Zinc Lozenges May Shorten the Duration of Colds: a Systematic Review

ID: 15728

Author: Dr Hemila

The manuscript was rejected on the basis of the following three reviewer comments.
However, I was encouraged to submit the manuscript to Clinical Pharmacology after taking into account the reviewer comments.

Replies to the reviewer comments by Harri Hemilä 24 Nov, 2010

The *reviewer comments* are in italics
Harri Hemilä's comments are followed after the bold **HH:**

REVIEWER 1 EVALUATION

There are a number of grammatical errors that need to be corrected before the paper can be accepted for publication.

HH: The language had been checked at the language center of our university, although I made some changes after the language checking.
In any case, reviewers 2 and 3 do not point out numerous grammatical errors so this is not an unanimous view by the 3 reviewers.
I would hope for some examples, what the reviewer means with the grammatical errors.

P. 5 "web of science, October 21, 2009", should read "web of science, October 21, 2010", this analysis should have taken place after the other searches not before.

HH: New searches were carried out.

Background, nature, scope and importance of the problem that led to the review/meta-analysis is well described.

It was not clear if the review/meta-analysis was guided by a written protocol, the paper was silent on this point, it should be addressed.

HH: No protocol was written. This was added to the Methods.

The aims/research questions could be more clearly described.

HH: Reviewer 1 writes: ”Background, nature, scope and importance of the problem … is well described” (above)
Reviewer 2 states: ”The introduction ... provides adequate … justification for the review.”
At the end of the Introduction I write ”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 zinc lozenges on the duration of colds in patients who had natural common cold infections.”
I do not understand what the reviewer would like to be changed.

The populations studied in which the results are to be generalised could be more specifically described, including a trial with children while all others involved adults diluted this aspect of the research and I note an analysis on adult only trials was not undertaken but would enable generalisability to this target population to be more clearer than the current mixing of population groups in the current presentation of results.

HH: The Macknin (1998) study - the only child study - is explicitly pointed out in the first paragraph of the results. It falls to the ”low-dose” studies. Thus, all ”high-dose” studies are ”adults only” studies. Thus, the finding that high-dose studies show substantial heterogeneity is based only on adult studies.

”in which the results are to be generalised” implies that the primary goal of the study is to calculate an estimate of zinc effect. In such a case it would be relevant to ask to whom does the estimate apply. However, the primary goal of this study was to find out whether the effect of zinc is heterogeneous over the dose. If it is, then there is no single estimate of effect that could be generalized.

The literature search was very basic, and raises the question if any trials were missed in the review. The time period for the searches was not provided, this needs to be inserted into the manuscript.

HH: The Cochrane collaboration uses Medline, Embase (close to Scopus, but I do not know the exact relation) and Cochrane Central data bases as universal requirement in searches. These three are used in this study. In addition, Web of Science is used as a different kind of approach. I do not see basis to argue that the search does not cover medical litarature widely enough.
Furthermore, statistically, I cannot see how any missing study could refute the heterogeneity seen between these 13 comparisons.
Reviewer 2 writes: ”The search strategy employed appears to be relevant and comprehensive.”

One trial yielded 3 comparisons so there were 11 trials providing data for 13 trials which potentially increases publication bias in reporting, but this was not addressed in the review. Also there is confused reporting of the number of trials vs comparisons, for example, page 11 states "13 zinc lozenge trials" when it should have stated "13 zinc lozenge comparisons", this needs addressing.

HH: The reviewer does not formulate any argument how the 3 zinc arms in one trial (Turner-2000) would lead to publication bias.
Publication bias means that researchers do not tend to publish negative studies, but publish all positive studies. All the Turner (2000) arms were ”negative” compared with the placebo arm. Thus, the Turner (2000) trial does not cause publication bias. Had it remained un-published, it would have

caused publication bias, but that question is not dependent on the number of arms in that trial.

I had been using term ”comparison” (because 1 trial had 3 comparisons). The correction noted by the reviewer was done.

The data extraction processes were well described. How missing data were handled was not described in the review.

HH: Imputation of missing SDs is described in Methods, and shown in the supplementary table.

The three high-dose Zn-acetate trials are a strong set indicating benefit. These are described ”... In the Petrus trial, only one participant was lost from follow-up [22]; in the first Prasad trial, two participants in the placebo group dropped out on day 2 [25], whereas there were no drop-outs in the second Prasad trial [21].” Thus, missing data for a few randomized participants cannot explain the findings of these 3 trials.

No measures were taken to identify or reduce the selective reporting of results within study reports or the selective reporting of studies themselves (publication bias), such as a funnel plot analysis.

HH: Selective reporting of results may be an issue, for example, in a pain-killer study if numerous different scales for pain are used, but only one (the most positive) outcome is reported.
The outcome analyzed in this meta-analysis is the duration of colds. There is no indication in the reports that any study used numerous definitions for the ”duration of cold” so that there could be reporting bias by selection only one of the outcomes.

”a funnel plot analysis”
"funnel plot" has been shown to be unsound e.g. in a paper published in BMJ:
Evidence based medicine: The case of the misleading funnel plot
http://www.bmj.com/content/333/7568/597
Therefore I do not see any point in using the funnel plot. I have never used it in any of my papers because I consider that it is unsound.

There seems to be a lack of reporting of the optional information size needed to detect the desired treatment effect.

HH: I do not understand what the reviewer means by ”optional information size needed to detect the desired treatment effect”.
In original trials, a power calculation should be carried out to find out whether the number of participants is adequate to test a quantitative hypothesis. I am familiar with numerous meta-analyses, but I have never seen a power calculation in them (power calculation might be reasonable in some special cases). I do not know whether I miss the point of the reviewer.

Statistical heterogeneity was examined and reported to be high for the overall results and the high dose subgroup but not for the low dose sub group. The author failed to interpret the high heterogeneity effect in terms of the results and generalisation, yet argued against previous research that demonstrated high heterogeneity. This needs to be addressed.

HH: I do not understand what this comment means. In the start of the Discussion I write:
”Dose-response relation between the quantity of zinc and the effect on the common cold duration: Trials where the total daily dose of zinc was less than 75 mg consistently failed to observe any benefit, whereas the majority of trials with higher zinc doses did (Tables 1 to 3, Fig. 1)...”

This means that ”negative” studies can be partly explained by low doses. The heterogeneity in high dose studies can partly be explained by the type of lozenge, as I describe later in the Discussion: ”Lozenge composition and the level of free zinc ions: Previously, several authors discussed the availability of zinc ions as a factor that may potentially modify the effect of zinc lozenges on the common cold [8-15]...” Thus, the meaning of the heterogeneity is discussed.

”...yet argued against previous research that demonstrated high heterogeneity.”
I do not understand what this means. I describe that ”... Jackson et al. [46,47] found statistically significant heterogeneity between the zinc trials... ” Heterogeneity means that all studies cannot be consistent with a uniform lack of effect. Consequently heterogeneity means that some people get benefit (or/and harm). However, Jackson et al. did not conclude that the kind of people or the kind of lozenges which are associated with a benefit should be searched for. They simply concluded that no effect has been proven (not correct because heterogeneity proves that there is effect on some people).
I refer to Thompson's paper in BMJ which proposes that, if there is heterogeneity, then the main focus of the reviewers should be on trying to understand the sources of the heterogeneity.

There was no discussion about the process of simply combining significance tests (P values) as the author did in this review. For example, studies with P values greater than 0.05 are published less often, making this method more prone to publication bias.

HH: In the Methods, I describe the reasoning for using the Fisher method. In the Discussion, I am not comparing the two methods (Fisher method and the forest plot) because the findings of the two methods are consistent. The Fisher method does not need imputation of SD values which is a strength of it, as pointed out in the Methods section.

”...For example, studies with P values greater than 0.05 are published less often, making this method more prone to publication bias.”
This is not a valid comment. If there is publication bias, that causes identical bias in the pooling of study results by the forest plot (not only in the Fisher method).

In terms of inclusion of papers, there was no mention if one or two reviewers undertook this process, no kappa value was reported in terms of inter-rater agreement.

HH: The criteria for including trials is described explicitly in Methods and there are no trials on which I had any hesitations when considering inclusion or exclusion.
I am the only author and I take full personal responsibility of the accuracy.

The results were placed in context and implications for the results presented but the paper needs further re-working before it can be accepted for publication.