Reviewer's report

Title: Efficacy of ambroxol lozenges for pharyngitis: a meta-analysis

Version: 3  Date: 20 November 2013

Reviewer: Borislav Dimitrov

Reviewer's report:

Results, p. 7 – Primary outcome

The main discussion point seems to be whether or not the calculated summary effect size (pooled mean difference) of -0.11 (95%CI -0.15 to -0.07) at 180 minutes for a dose of 20mg Ambroxol vs placebo, is of clinical relevance (clinically important). As of note, on a scale of -1.0, this represents a difference (reduction) of -11% that may go up to -15% (and, in some trials, from -2% up to -26%).

However, any other or further representations of this effect size are derivatives (e.g., standardised differences, probability indexes, etc.). Certainly, together with the pooled estimates from the meta-analysis, they represent a clear addition to knowledge being provided by this review as compared to the previous one.

Some of these derivatives may be standardised (adjusted) and some may take into account arbitrary cut-offs (thresholds) but all are secondary to the primary measure and rely on additional methods and assumptions. Whether or not these additional secondary representations give further evidence on the clinical significance of the primary measure of -0.11, is a matter of interpretation rather than a methodological issue. For instance, if I interpret correctly the standardised mean difference of -0.34 (20mg) on a scale of maximum -1.0, this would say that the mean effect size here would be even -34% (and may go up to -46%) – about one third of a mean difference, if true, may not be easily neglected. In the same time, the probability of 59% (20mg), as obtained through the probabilistic index (as newly developed and published in 2012 by one of the Authors) may indicate an effect that is not too much different from 50% (or, 0.5) as probably used as a cut-off. Of course, all of the above conjectures are true if I have understood correctly the methodological approach of the Authors (if not, maybe I am wrong, or the Authors may need to further describe or present their approach and results in a more detail).

The fact that both reviews provide almost equal estimates of these effects sizes in each of the 5 included trials (including 95%CIs), is of confirmatory nature, as they used exactly the same methods and computational techniques, including the fixed-effects model approach. In this sense, I would accept the pooled estimate since the heterogeneity is very small, virtually zero (please, can the Authors check this nil again!) and the test of overall effect is highly statistically significant.

Clearly, this difference of -0.11 (as a pooled estimate) is statistically significant as a reduction of pain. Whether or not it is clinically relevant, as a patient-reported
outcome (PRO), I am not in a position to judge. Usually, in psychometrics, it is referred to as “minimal clinically important difference” (MCID). One of the Authors has reviewed in his 2012’s paper various approaches to clinical importance (relevance), but if useful, the Authors may wish also to consider, in this sense, the reference by Crosby (http://www.rygforskning.dk/sites/default/files/files/articles/crosby.pdf) – there maybe another way of assessing the clinical significance of the statistically significant findings.

On the other hand, methodologically, the fact that the Authors and the previous review (Ref.14) use information from the same 5 trials and use the same computational techniques does not make automatically the primary effect size measure a true measure of the real effect.

In particular, the primary effect size for each trial is calculated as an Area Under the Curve (AUC) which is a better summary statistic than the simple, overall mean value (a reference is attached as a PDF file) when repeated measures at unequal intervals are included in the calculations, as in this case (particularly, when the data are of “peaked” type).

However, while I agree with the AUC as a summary statistic in this case, per se, I have some concerns over the equation used to produce the AUC as an arithmetic estimate of the true mean value (including the algorithm to divide over 3, as described by both reviews). Interestingly, no one of the two reviews provides a reference (source) for this equation (or, at least, I was not able to find one). I am not saying that the equation is wrong; what I am saying is that I cannot trace its source from the two reviews (neither any other reader would!). It may be a common sense equation for calculation of AUC, but I would still suggest that this equation needs a reference. For instance, the reference I attach also contains an equation for calculation of AUC, not only as an arithmetic mean, but also as a geometric mean.

Given the inconsistencies, uncertainties (e.g., external validity) and the lack of detail or references in both reviews, I would suggest the following to the Authors of this manuscript that might strengthen the evidence base of what they present as a measure of effect size at 20mg Ambroxol:

- Minor Essential Revisions:
  1) Please provide reference for the equation that you have used to calculate the AUC;
  2) You assumed a normal distribution (see Methods, p.5, last paragraph); however, your argument may not hold true – have you used any formal tests to assess normality of distributions;
  3) I do not have the individual values at each time point for each trial, but another way of assessing normality would be if the standard deviations change with the mean values. If not normally distributed, I would suggest to use a logarithmically transformed values or other ways to achieve (possibly) a normal distribution – thus you may be able to compute a better, more precise estimates of the true effect sizes;
4) I would suggest that if available, you can prepare and publish in a separate table (may be as an online Appendix only) the individual value at each time point for each trial on which you have based your calculation of the AUC for 20 mg Ambroxol (should be probably 5 trials x 2 arms x 5 time points = 50 values);

5) In parallel, if arithmetic means are normally distributed or not, you may still calculate and use the geometric means of the AUCs (see attached reference PDF); thus, also a formal test for statistical significance for the ratio of the geometric means may be also applied as an estimate of the difference between the two arms, if needed;

6) You used a fixed-effect model approach. Can you use also a random-effect model approach to see how the estimates/related 95% CIs may change? At least, this may give additional viewpoint on the statistical significance of the findings and their meaning;

7) You assume a publication bias; can you provide a funnel plot and/or some formal statistical tests for this conclusion?

I would also point out, given the above discrepant aspects, that a cost-effectiveness analysis (CEA) may provide further evidence about the clinical relevance (practical usefulness) of the computed estimates on a wider scale. However, I understand that this is well beyond the scope of the current paper. If CEA is not available at this point, then I would suggest that we, as Authors and reviewers, do our best to provide best possible available evidence in this paper and leave the clinical judgement on the clinical significance of the statistically significant effects to the readers.

I think that the authors can be trusted to make these additional clarifications and additions. I hope that this will make the manuscript better and further improve its quality as an integrated piece of work.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I do not have any competing interests