

Vitamin C and exercise-induced bronchoconstriction: further problems in the Cochrane review “vitamin C for asthma” (2012)

Harri Hemilä

Department of Public Health
University of Helsinki
Finland

<http://www.mv.helsinki.fi/home/hemila>

harri.hemila@helsinki.fi

Version 10 Sept, 2013

Background

In **2009**, I submitted Feedback about the Cochrane review:

**Kaur B, Rowe BH, Arnold E. Vitamin C supplementation for asthma.
Cochrane Database Syst Rev. 2009 Jan 21;(1):CD000993.**

<http://www.ncbi.nlm.nih.gov/pubmed/19160185>

<http://dx.doi.org/10.1002/14651858.CD000993.pub3>

The 2009 feedback is available within the Cochrane review (link above) and as a separate document at:

<http://hdl.handle.net/10138/38500>

and

<http://www.mv.helsinki.fi/home/hemila/H35P.pdf>

The **2009** feedback had two components.

First, it documented large errors in the extraction of data, and also with data analysis, in 3 RCTs on exercise-induced bronchoconstriction [EIB].

Second, it presented a brief meta-analysis of 3 RCTs on EIB. I calculated that the 3 RCTs on EIB give strong evidence that vitamin C alleviates EIB with a combined $P = 0.00007$.

The second component of the 2009 feedback was extended to a full-length meta-analysis and was published in June 2013:

<http://www.ncbi.nlm.nih.gov/pubmed/23794586>

<http://dx.doi.org/10.1136/bmjopen-2012-002416>

Three years after the **2009** feedback, in **2012** the Cochrane review authors replied to my criticism. The authors' 2012 response is shown at page 3.

This document, starting at page 4, deals with the problems with the authors' 2012 reply to my 2009 criticism.

Background

On 10 June 2013, the comment starting at page 4 was submitted to Cochrane Airways group. On 13 June 2013, it was rejected with the following argument:

“It is not the policy of the Cochrane Airways Group for the feedback section of Cochrane Reviews to serve as a forum for commentators to debate whether they agree with the way in which the authors of the review have responded to their previous comment.

We will not be publishing your latest comment in the review of vitamin C for asthma.”

Yours sincerely, Toby Lasserson [Feedback Editor with the Cochrane Airways Group]

This rejection is inconsistent with the Principles of the Cochrane Collaboration:
<http://www.cochrane.org/about-us/our-principles> (accessed 10 Sept 2013)

Eight principle is about “**Ensuring quality**”. It is formulated as follows:

“Ensuring quality: by being open and responsive to criticism, applying advances in methodology, and developing systems for quality improvement”

The policy by the Feedback Editor is inconsistent with being “open and responsive to criticism”. Argument “... not ... to serve as a forum for commentators to debate whether they agree with the way in which the authors of the review have responded to their previous comment” prevents scientific discussion, and does not keep the Cochrane review authors accountable for the validity of their arguments and errors in their reviews.

Schriger and Altman commented on the importance of post-publication review of research:
<http://www.bmj.com/content/341/bmj.c3803>

“... Post-publication critique, ... is a sign of a healthy scientific community, a community actively working to move the field forward. ... so an important flaw in a publication should be highlighted whenever identified. It is therefore regrettable that journals discourage letters through needless constraints ... their refusal to publish several rounds of letters. ... journals have a strong conflict of interest regarding letters to the editor because publishing criticisms of journal articles suggests that the editors are not doing their job and may lower the prestige of the journal... The inadequacy of authors’ responses to criticism suggests that authors feel no obligation to respond to reasoned criticism...”

There is an obvious conflict of interest problem when Cochrane Airways group declines to publish a second feedback which points out errors in their review. With such policy, the reviews of the Airways group appear better than they actually are.

The Cochrane review on vitamin C and asthma was first published in 2001 and the errors pointed out in the 2009 feedback originated from the first version:

Kaur B, Rowe BH, Ram FS. Vitamin C supplementation for asthma.

Cochrane Database Syst Rev **2001**;(4):CD000993.

<http://www.ncbi.nlm.nih.gov/pubmed/11687089>

<http://dx.doi.org/10.1002/14651858.CD000993>

Thus, the Cochrane review has been misleading readers for a decade, and is still misleading the readers (September 2013) even after the 2009 feedback had pointed out serious errors.

**Vitamin C and exercise-induced bronchoconstriction:
further problems in the Cochrane review “vitamin C for asthma” (2012)**

The Cochrane review authors wrote the following reply to my 2009 criticism.

Their reply is published in:

<http://dx.doi.org/10.1002/14651858.CD000993.pub3>

“Reply:

This comment on the trials relating to exercise induced bronchoconstriction (EIB) was submitted in March 2009 and published alongside the review in November 2010.

We thank Dr Hemilä for the feedback, but do not think that the technical issues raised over the analysis of data from the three small cross-over trials (including a total of 40 participants), substantively alter the strength or direction of the results, the quality of the evidence, or the conclusions of the review.

We agree that crossover trials are best analysed using paired t-tests, but do not agree with the presentation of one-tail P values above.

A two-tailed paired t-test did not show a statistically significant difference in change in FEV1 either immediately after exercise (shown in analysis 1.2) or five minutes later in Schachter 1982 ($P = 0.18$ and 0.057 from Table II and Table III respectively). Therefore the author’s choice not to include the latter observation does not mislead the reader in our opinion.

We agree that the mean differences in FEV1 reported from only 11 of the 20 participants in Cohen 1997 should not be included in the review, and this has been removed from the analyses.

The authors entered data for the fall in FEV1 from Tecklenburg 2007, using a standard error derived from a conservative estimate of the P value based on the paired t-test (reported in the paper as $P < 0.05$). We see no compelling reason to overturn this approach since the average effect is unaltered and the data come from a study of only eight participants.

We agree that the baseline lung function is not a useful outcome for this review and have removed the pre-exercise outcomes.

We do not agree with the suggested approach of combining P values from Cohen 1997, Schachter 1982 and Tecklenburg 2007 in view of the clinical heterogeneity between the studies and outcomes under consideration. Such an approach focuses attention on whether any effect observed is attributable to chance. This is itself potentially misleading since it does not take account of the magnitude of effect across the studies. The analyses presented in the review have now been amended so that only mean differences and confidence intervals for the studies are presented, and not associated P values.

We are content for readers to consider the comment from Dr Hemilä alongside our response, and to make up their own minds regarding the authors’ approach to the analysis of data and the conclusions of the review.

Posted by Emma Welsh, Managing Editor of the Cochrane Airways Group, on behalf of the author and editorial teams.”

**Vitamin C and exercise-induced bronchoconstriction:
further problems in the Cochrane review “vitamin C for asthma” (2012)**

Harri Hemilä

“We ... do not think that the technical issues raised over the analysis of data from the three small cross-over trials (including a total of 40 participants), substantively alter the strength or direction of the results, the quality of the evidence, or the conclusions of the review.”

HH: The number of observations that are needed to show an effect depends on the magnitude of the effect. If the effect is small, a great number, in some cases tens of thousands of participants will be needed. However, if the effect is large, then only a small number of participants will be sufficient to show an effect. Furthermore, if the effect is very large, no trials will be needed (1).

When a small study *does not find* a difference between intervention and control, the result may be a false negative because of low statistical power of the study. However, when a small study *does find* a statistically highly significant effect, it cannot be ignored with similar reasoning. Therefore, the Cochrane review authors' argument is not valid. Three RCTs on exercise-induced bronchoconstriction [EIB] (2-4) obtained similar results with no evidence of heterogeneity between the trials. They should not be ignored on the basis of the low number of participants.

Furthermore, asthma is a heterogeneous syndrome, which comprises a collection of different phenotypes rather than a single disease (5). Different asthma phenotypes should be considered separately. Thus, the results of all asthma-related studies should not be pooled into a uniform “**direction of the results**” kind of conclusion, instead the EIB studies should be analysed as a distinct phenotype.

“We agree that crossover trials are best analysed using paired t-tests, but do not agree with the presentation of one-tail P values above.”

HH: It is good that the authors agree that cross-over studies should not be analyzed by unpaired t-tests. However, the error was the same in the first version (2001) of the Cochrane review (6) and thus this error has been misleading readers for a decade. Furthermore, the same error still remains in Analysis 1.1 of the 2012 version in which the FEV₁ data of Schachter 1982 table II (2) are shown as un-paired data, although the data are paired.

The comment on the 1-tailed P-values is not sound. P-values must be 1-tailed for the Fisher method of combining P-values (7). Furthermore, the transformation from P[1-tail] to P[2-tail] is a trivial multiplication, if a reader prefers the 2-tailed values. I calculated in the 2009 Feedback that “P[1-t] = 0.0005” for the Cohen 1997 trial on EIB (3), which gives a P[2-t] = 0.0010 with a simple calculation. Such a low P[1-t or 2-t] value indicates that it is highly unlikely that the difference between the vitamin C and placebo days is caused by chance alone.

There is nothing wrong with using 1-tailed P-values, but the tails must be stated explicitly in a study methodology. For example, one major textbook of medical statistics writes: “*the distinction between one- or two-sided p-values is not of fundamental importance in interpreting the results of a study, provided that the investigator clearly states which type of statistic was used*” (8, p. 183) and “*From a statistical point of view, the report based on a two-sided p-value should be 'alike' (not significantly different) or 'different', but not that 'the innovation is more effective than standard therapy'. Both investigators and readers might be surprised to learn that a study has shown only that two treatments are different, without commenting on directionality, but this is the proper interpretation of the two-sided test*” (8, p. 184).

The response “*we do not agree*” is not a scientific argument. The authors should explain why they do not agree and how their disagreement influences the conclusions. As stated above, the transformation from P[1-t] to P[2-t] is trivial arithmetic.

“A two-tailed paired t-test did not show a statistically significant difference in change in FEV1 either immediately after exercise (shown in analysis 1.2) or five minutes later in Schachter 1982 (P = 0.18 and 0.057 from Table II and Table III respectively).”

HH: In my 2009 Feedback, I wrote that “*In EIB the fall in FEV1 occurs 5 to 20 minutes after the end of exercise (Rundell 2009 [ref. 9]), and even Schachter reported that, on the screening day, there was no fall in FEV1 immediately after exercise, but a significant fall 5 minutes after the exercise (Schachter 1982 Fig. 2 [ref. 2]).*”

Why have the authors ignored this comment?

The authors do not describe any rationale to keep the “*immediately after exercise*” data in the current version, although it is irrelevant for EIB. Furthermore, Analysis 1.2 (mentioned above) presents FVC data (0 min data) and not the FEV₁ data, thus the statement above is sloppy. The FEV₁ data (0 min data) is shown in Analysis 1.1.

Why do not the authors present the “five minutes later” data for FEV₁ in the Cochrane review (2012) although the 5 min data is relevant whereas the 0 min data is not [2,9]?

“We agree that the mean differences in FEV1 reported from only 11 of the 20 participants in Cohen 1997 should not be included in the review, and this has been removed from the analyses.”

HH: This is not appropriate. The correction of one error has led to a new error. The previously presented subgroup of 11 participants was biased because it was selected on the basis of the benefit (3). However, there is no justification to exclude the Cohen 1997 trial. Cohen 1997 included

20 participants with EIB (3). On the placebo day, all 20 participants suffered from EIB. On the vitamin C day, 10 participants suffered from EIB but 10 participants did not. There is no missing data problem with the dichotomous EIB outcome.

Thus, the Cohen 1997 trial gives a vitamin C effect estimate that the proportion of people who suffer from EIB can be reduced by 50 percentage points. As described above, the difference between placebo and vitamin C corresponds to P[2-t] = 0.001 and thus the 50 percentage point difference cannot be easily explained by random variation. In the 2012 version, the Cochrane review authors do not give any reason for ignoring the effect of vitamin C in this methodologically sound RCT. The authors should have calculated the estimate of vitamin C effect and its 95% CI.

“We agree that the baseline lung function is not a useful outcome for this review and have removed the pre-exercise outcomes.”

HH: It is good that the authors agree that a baseline measurement is not an outcome. However, the calculations were the same in the first version (2001) of the Cochrane review (6) and these calculations had been misleading readers for a decade.

“We do not agree with the suggested approach of combining P values from Cohen 1997, Schachter 1982 and Tecklenburg 2007 in view of the clinical heterogeneity between the studies and outcomes under consideration. Such an approach focuses attention on whether any effect observed is attributable to chance. This is itself potentially misleading since it does not take account of the magnitude of effect across the studies. The analyses presented in the review have now been amended so that only mean differences and confidence intervals for the studies are presented, and not associated P values.”

HH: This is not a reasonable argument. There are two usual goals in medical statistics:

- 1) to find out whether the difference between intervention and control can be explained by chance (i.e. calculation of the P-value) and
- 2) to estimate the magnitude of the intervention effect.

The confidence interval approach gives information on both objectives, but both can also be approached separately. In the 2009 Feedback, I calculated that the three RCTs give a combined $P = 0.00007$. The benefit of vitamin C is not explained by chance. That is answer to goal #1.

Answer to goal #2 can be obtained by looking at the findings. Cohen found a **50%** decrease in the occurrence of EIB (see above, ref. 3). Schachter found a **45%** decrease in the FEV₁ decline 5 min after exercise (Table III: [0.44-0.24]/0.44, ref. 2). Tecklenburg found a **50%** decrease in the FEV₁ decline ([12.9-6.4]/12.9, ref. 4). Thus, all three RCTs found about a 50% effect from vitamin C administration. Thus, the **“magnitude of effect across studies”** can be considered in addition to the combined P-value.

The authors write: **“the clinical heterogeneity between the studies and outcomes under consideration.”** (see above)

HH: This is not a valid comment. All of the three RCTs exposed their participants to a short term physical stress and all the studies measured FEV₁ before and after the exercise challenge test (2-4). Thus the exposure and outcome are not heterogeneous. There are indeed differences between the three RCTs, but these differences make the generalizations about the results wider, rather than precluding a meta-analysis. The three RCTs were carried out in three different decades and on two different continents. In addition, the mean age was 14 yr in the Cohen study, but 25 and 26 yr in the two other studies, and also the criteria for EIB in the studies differed. Nevertheless, all of these studies found a 50% reduction in the FEV₁ decline caused by exercise. Thus, the heterogeneity between the studies makes the generalization of the 50% effect to be much wider compared with three trials with very similar participants and settings.

It is obvious that a short comment cannot include a detailed statistical analysis of the three RCTs on vitamin C administration and EIB, and that was not the purpose of my 2009 Feedback. Instead the primary purpose was to point out the severe shortcomings in the Cochrane review, which originated in 2001 with Balvinder Kaur, Brian Rowe and Felix Ram as the authors (6).

The 2009 brief analysis was extended to a full length meta-analysis (10). On the basis of the three RCTs on EIB (2-4): **“The pooled relative effect estimate indicated a 48% reduction (95%CI: 33% to 64%) in the post-exercise FEV₁ decline when vitamin C was administered before exercise.”**

That pooled 95%CI corresponds to $P(2-t) = 0.0000000001$ ($Z = 6.3$; Fig. 4 in ref. 10). Thus, the combined P-value which I calculated in the short 2009 Feedback ($P = 0.00007$) is not an exaggeration, indeed it is an underestimate of the statistical significance.

Since two of the trials on vitamin C and EIB were published in 1980s [2] and 1990s [3], the authors of the Cochrane review could have concluded already in 2001 (6) that there was strong evidence to suggest that vitamin C administration decreases the exercise-induced decline in FEV₁.

References

1. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007;334:349-51.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1800999>
2. Schachter EN, Schlesinger A. The attenuation of exercise-induced bronchospasm by ascorbic acid. *Ann Allergy* 1982;49:146-51.
<http://www.ncbi.nlm.nih.gov/pubmed/7114587>
http://www.mv.helsinki.fi/home/hemila/A/Schachter_1982_ch.pdf
3. Cohen HA, Neuman I, Nahum H. Blocking effect of vitamin C in exercise-induced asthma. *Arch Pediatr Adolescent Med* 1997;151:367-70.
<http://dx.doi.org/10.1001/archpedi.1997.02170410041005>
4. Tecklenburg SL, Mickleborough TD, Fly AD, et al. Ascorbic acid supplementation attenuates exercise-induced bronchoconstriction in patients with asthma. *Respir Med* 2007;101:1770-8.
<http://dx.doi.org/10.1016/j.rmed.2007.02.014>
5. Wenzel SE. Asthma: defining of the persistent adult phenotypes. *Lancet* 2006;368:804-13.
<http://www.ncbi.nlm.nih.gov/pubmed/16935691>
6. Kaur B, Rowe BH, Ram FS. Vitamin C supplementation for asthma. *Cochrane Database Syst Rev* 2001;(4):CD000993.
<http://www.ncbi.nlm.nih.gov/pubmed/11687089>
<http://dx.doi.org/10.1002/14651858.CD000993>
7. Fisher RA. Combining independent tests of significance. *Am Stat* 1948;2:30. See a summary of the method:
http://en.wikipedia.org/wiki/Fisher's_method
8. Bailar JC, Hoaglin DC (eds). *Medical Uses of Statistics*, 3rd ed. N Engl J Med and Wiley, New Jersey, 2009.
9. Rundell KW, Slee JB. Exercise and other indirect challenges to demonstrate asthma or exercise-induced bronchoconstriction in athletes. *J Allergy Clin Immunol* 2008;122:238-46.
<http://www.ncbi.nlm.nih.gov/pubmed/18678339>
10. Hemilä H. Vitamin C may alleviate exercise-induced bronchoconstriction: a meta-analysis. *BMJ Open* 2013;3:e002416.
<http://www.ncbi.nlm.nih.gov/pubmed/23794586>
<http://dx.doi.org/10.1136/bmjopen-2012-002416>