Author's response to reviews

**Title:** Inhaled Drugs to Reduce Exacerbations in COPD Patients: a Network Meta-Analysis

**Authors:**

Milo A Puhan (milo.puhan@usz.ch)
Lucas M Bachmann (lucas.bachmann@usz.ch)
Jos Kleijnen (jos@systematic-reviews.com)
Gerben ter Riet (G.terRiet@amc.uva.nl)
Alphons G Kessles (akes@aozg.azm.nl)

**Version:** 2 **Date:** 16 November 2008

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

We thank you and the reviewers’ very valuable evaluations and for your invitation to revise our manuscript. Please find below our answers the comments. We restated each comment followed by our response in italics. We have submitted two versions of the revised manuscript, one with tracked changes and one where all changes are accepted.

Reviewer 1

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

1. Sources: EMBASE, regulatory authority websites in the US and UK, and manufacturers trial registries were not mentioned in data sources and selection. These sources are important to be considered, indeed. Since a number of systematic reviews on inhaled drugs for COPD patients were published already, among those some Cochrane reviews of high quality, we relied on the searches of these existing systematic reviews in order to avoid unnecessary duplication of previous work. In these systematic reviews, EMBASE and regulatory authority websites were searched for trials and drug companies were asked for unpublished studies. In the revised paper we provide more details about the searches of these existing reviews: “We based our searches on existing systematic reviews in order to avoid unnecessary duplication of previous work. The existing systematic reviews used extensive search strategies that included several databases such as Medline, EMBASE, CINAHL and LILACS as well as websites of regulatory bodies. In addition, drug companies were approached for unpublished trials.”

2. Why the authors did not include mortality as well as exacerbation rate in the analysis? There were enough studies (18?) with at least 6 months to dismiss mortality rate as the major outcome. Furthermore, there are at least 2 reviews with opposite findings regarding mortality and these reviews included studies with # 4 weeks duration. It would be good to have mortality as an outcome and we acknowledged in the discussion that our focus on exacerbations is a limitation, indeed. However, when we planned this analysis little evidence on the effects of inhaled drugs on mortality was available and still is. We added this explanation to the revised manuscript: “When we planned the current analyses the body of evidence on mortality was too small to provide,
even if pooled, precise effect estimates. The TORCH trial[14] published in 2007 was the first trial that was powered to assess the effects of inhaled drugs on mortality, but even this very large trial turned out to be too small. Therefore, we decided to postpone analyses on mortality to a point in time where more data would become available.”

We discuss in the discussion section that pooling of reported data on exacerbation rates is difficult in meta-analyses. If all available trials were analyzed and reported adequately a pooled analysis would be possible. But as Suissa et al pointed out, only few trials took in their (Poisson regression) analyses within- and between-patient variability in exacerbation rates into account so that published data are not very reliable. We are aware that meta-analysts treat such exacerbation rate data as continuous data (mean rates per year and SD) and pool them (by calculating weighted mean differences) but we think that these analyses do not take into account the complicated structure of these data. The revised text reads as:

“A limitation of any approach to pool data on exacerbations is that (relative) treatment effects can only be estimated adequately based on the proportion of patients with at least one exacerbation. Analysis of exacerbation rates expressed as mean exacerbation rates per person-year would offer a more comprehensive use of the data and it would be less dependent on the length of follow-up. But such data cannot be pooled adequately without having access to adequate and fully reported analyses that took into account within- and between-patient variability in exacerbations. But as Suissa pointed out most trials were not analyzed and reported adequately.[19]. Ideally, an individual patient data meta-analysis would be conducted but it is very challenging to convince investigators of all relevant trials to share their data[64].”

3. Ref 53. The mean age of 44.8 yr in the placebo arm, the great proportion of never smokers and the mean FEV1 of 83.3 and 86 % for fluticasone and placebo respectively, confirmed the old tendency of Dutch authors to include both asthmatic and COPD subjects. Then, this study should not be included because FEV1 post bronchodilator must be less than 80% pred. (COPD definition according to GOLD). DIMCA means Detection, Intervention and Monitoring of COPD and Asthma program.

We agree that the patient population was younger and had less airflow obstruction that the populations of other trials. And we had discussion about whether to in- or exclude this trial in the group as well. However, we think that we have to include the trial nevertheless. Mean FEV1/FVC was <0.7, which is together with an FEV1 >80%, in
accordance with a diagnosis of COPD GOLD stage I. Therefore, we find it hard to justify excluding the trial.

COPD is a very heterogeneous disease and when doing systematic reviews it is often challenging to decide whether all or the majority of patients had COPD reporting of clinical and spirometric characteristics is often poor. In the revised manuscript we discuss this common problem as follows: “A common problem of systematic reviews in COPD is poor reporting of clinical and spirometric characteristics of patients enrolled in included trials. It is sometimes difficult to judge whether all patients had COPD or if some patients had other lung disease such as asthma. The trial by van der Boom et al [53] is an example. Also, there is often no separate reporting for moderate and severe exacerbations although it would be informative to estimate the effects of inhaled steroids stratified for the severity of exacerbations. COPD is a heterogeneous disease and it is important that future studies provide more information about their patients including clinical characteristics, lung function data but also information on co-treatments.”

4. Page 16. Since mortality was not taken into account in this review; it would be too speculative to recommend treatments.
We agree that the decision for or against inhaled steroids should not be based on the effects on exacerbations only. We added a paragraph about balancing benefits and adverse events (see further below) and tempered our conclusion, which now reads as: “The results of our study may support physicians in selecting inhaled drug treatments for COPD patients. In general, long-acting beta-agonists or anticholinergics appear to be the treatment of choice to reduce exacerbation rates. Adding an inhaled corticosteroid does not provide additional protection from exacerbations. In patients with low FEV₁, combination treatment and long-acting anticholinergics should be favored because they reduce the risk for exacerbations more than single treatment with a long-acting beta-agonist. Our analyses together with the existing meta-analyses that considered additional outcomes inform patients and physicians to balance benefits and downsides of different inhaled drug treatments for COPD.”

5. Could the authors find any difference if only considering hospitalization or systemic corticosteroids as criteria of COPD exacerbation?
It is an important question to learn more about the effects of inhaled drug treatments on moderate or severe exacerbations. However, the distinction between moderate and
severe exacerbations was not reported by many trials so that we could not address this issue. We performed, nevertheless, an analysis stratified for the definition of exacerbations (symptom- or event based) but could not find a significant influence. As described above (reply to comment 3) we added a paragraph about the problem of poor reporting.

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. Appendix 2 numbers between parentheses did not correspond to the reference number.
   We thank the reviewer for this comment and corrected the reference numbers.

2. Fig 2 and 3 did not show significance (p) neither their legends.
   We do not think that p-values provide additional value. We reported point estimates and confidence intervals that allow a quantitative appreciation of the results. We prefer not to add unnecessary p values to the figures in order not to overload them. But if the editors and reviewers feel that we should add them we could still do so.

Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

1. Page 4:....the predominant question in clinical practice is to choose between treatments rather than deciding whether to treat or not to treat[7, 8]. I think that subjects with FEV1 < 50% predicted, are treated by most of physicians with triple-drugs (ICS+LABA+LAAC) such as in the INSPIRE study (Thorax 2008).
   We agree with the reviewer that many physicians prescribe triple-therapy though precise numbers about this prescription pattern are scarce yet. We are not fully clear to which trial the reviewer refers. The INSPIRE trial led by Prof Wedzicha was published in the American J of Resp and Crit Care Med and randomised patients to long-acting bronchodilators plus inhaled steroids or to tiotropium. We have included this trial in our analysis. We found a trial investigating triple therapy (Aaron et al Ann Intern Med 2007, 146(8):545-555.) but we did not include that trial in our analysis because the treatments combinations were not comparable to those of the other trials. However, as soon as
more trials abut triple therapy become available we will be happy to include them in an updated analysis.

2. The authors could estimate the fail-safe numbers (N) by the Rosenberg’s method to assess for the influence of publication bias on the meta-analysis. (Rosenberg MS. The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. Evolution Int J Org Evolution. 2005;59(2):464-468.) The fail safe N, represents the number of undetected negative studies that would be needed to change the conclusions of a meta-analysis. There is one more article: Gleser LJ, Olkin I. Models for estimating the number of unpublished studies. Stat Med 1996; 15:2493–2507.

The reviewer makes an interesting suggestion about an additional analysis. Since our analysis is quite comprehensive already we would like not to add it to our paper. We also presume some publication bias may play a major role in COPD trials. Investigating publication is a major endeavor that would go beyond the scope of this paper because it would need to consider additional outcomes and to address the problem of selective outcome reporting.

3. Could the authors calculate the number needed to treat (NNT) with every treatment arm in order to prevent a severe exacerbation when comparing with placebo?

This is a valuable suggestion since severe exacerbations are those bearing heavily on the patients’ quality of life and on prognosis. But there are several problems to calculate NNTs. As explained above most trials did not report on moderate and severe exacerbations separately. Also, although it is possible to calculate NNT by calculating attributed risks and by assuming a common risk for exacerbations across control groups. But risks for exacerbations varied much across trials so that we do not think that such an assumption would be sensible. Also, we would also like to emphasize that our main comparisons were not those with placebo but the comparisons among active treatments. Finally, the number needed to treat refers to a specific time period (e.g. one year) but there was substantial variability in observation time across included trials so that we are uncertain whether referring to a specific treatment duration such as for example six months would be justifiable. Finally, interpretation of numbers needed to treat is most meaningful if all relevant outcomes (including harms) are considered so that balanced
decision can be taken for or against a treatment. Therefore, we would like not to calculated numbers needed to treat in this paper if the editors and reviewers agree.

4. The cut off point < 40% pred FEV1 was important; because GOLD recommended ICS under 50% plus exacerbations in symptomatic subjects. It was interesting that there was no effect modification when stratified for FEV1% or # 50%. This evidence could induce a modification in the next updating of GOLD.
   We thank the reviewer for this comment. Current cut-offs are based on consensus rather than on empirical data. We hope that other studies also try different cut-offs in order to find out what cut-offs have prognostic value.

5. Page 15. I am not sure if physicians are more familiar with odds ratios or relative risks than with mean differences in exacerbation rates or worthy percentage of reduction with any treatment.
   We agree that this is actually uncertain. Therefore, we are more cautious in the revised the paper, which reads as: “Also, physicians may be more familiar with this format and respective estimates for treatment effects such as odds ratios or relative risks than with mean differences in exacerbation rates. However, there is little evidence on how effect estimates should be presented in order to facilitate the transfer from research into practice.”

6. The authors wrote that their review could help physicians in indicating treatment. However, I wonder which combined therapy and which ICS dose? There were lesser publications with formoterol/budesonide than with salmeterol/fluticasone and different doses of fluticasone and budesonide alone. We (physicians) need more consistent help for the day to day practice.
   We agree that there is very little guidance on this. However, given the rather small effects of these drugs (see comparisons with placebo in figure 2) it is difficult to detect differences between them and between different doses. Another challenge is that comparisons between drugs are difficult because doses varied across trials. When considering all variables that influence between-drug comparisons strata of trials become so small that even in meta-analysis it is hardly possible to detect any differences. Theoretically, it would be possible to investigate the effects of different
doses with our analytic method but the data on inhaled drugs in COPD are not suitable to do this.

7. Ongoing trials with LABA and tiotropium from a single device probably confirm the hypothesis that it was largely better not to use single therapy in COPD despite severity. This comment closely relates to comment 1. We agree that treatment with two types of bronchodilators (± inhaled steroid) probably provides the greatest effects but evidence on this is still evolving. We expect that in a few years more trials will be published so that that triple treatment can be considered in an updated analysis. Also, while most single trials are too small to study effect modification (e.g. by disease severity) a pooled analysis like the one we present in the manuscript is attractive because based on a much larger data set.

Reviewer 2
This is a well written and presented manuscript. I read the manuscript with great interest as it provides an alternative method for systematic reviewing of RCT’s and data analysis. The methodology looks sound and the results well presented. Particularly, interesting to read the results for patients with FEV1 < 40%. However, two main issues require consideration. (1) The study results suggest that ICS are of limited benefit in COPD. This does go against normal recommendations and COPD pathology. We are aware that ICs usual takes much longer to have a clinical effect (particularly on exacerbations) and the authors did not look at time as a factor in their analysis. The studies included vary from 3-36 months. It is important that the authors do analysis based on time. For example studies that were conducted for longer than 12 months duration of which there are 11 included in the review. It would be interesting to see these results.

We thank the reviewer for his comment and interest in our analytic method. We agree with the reviewer that our results go against normal recommendations. Today’s COPD recommendations, however, are only partly based on evidence. For example, the preeminent GOLD guidelines are not based on systematic reviews of the literature and the drug industry was heavily involved in their development. The rather poor quality of existing COPD guidelines triggered a major initiative by the ERS, ATS, WHO and many other organizations that aims at developing evidence-based guidelines for COPD
Whether our results go against pathology is controversial. While some argue that inhaled steroids reduce inflammation in the airways, there are others who argue that the inflammation in COPD is, at least sometimes, not suppressed by corticosteroids but rather worsened (see for example Barnes Europ Resp Journal 2004).

The reviewer raises the important point of the length of treatment. In our manuscript, we already described an analysis stratified for the length of treatment (≤ and >6 months) that did not suggest that the length of treatment had an influence on effect estimates. We conducted another analysis with stratification for ≤ and >12 months but again, we did not find any difference between trials with a treatment duration of ≤12 months (compared with placebo, odds ratios for long-acting beta-agonists 0.77, long-acting anticholinergic 0.72, inhaled corticosteroids 0.71 and combined treatment 0.62) and >12 months compared with placebo, odds ratios for long-acting beta-agonists 0.75, long-acting anticholinergic 0.68, inhaled corticosteroids 0.81 and combined treatment 0.77). We report this additional analysis in the revised paper.

(2) The search for studies were last done in November 2007, this requires updating.

We agree that our searches are not as of November 2008 but we think that it is too early to update the search. There is increasing agreement that a time-based approach with updates every two years is often not an efficient approach considering the small proportion of updated reviews that result in a changed conclusion. In our example, there are so many trials and patients included that an update is unlikely to change the findings unless a number of very large trials show results that differ substantially from the current analyses. But even with a time-based approach time our search is not out-of-date since our search or those of the existing systematic reviews, respectively, was still performed considerably less than two years ago. If the editors and reviewers agree we would like not to update the search at this moment.

Reviewer 3

This is a systematic review and network meta-analysis to assess the relative effects of common inhaled drugs—long-acting beta-agonists (LABAs), long-acting anticholinergics, inhaled corticosteroids (ICS), LABA+ICS—to reduce exacerbations in patients with chronic obstructive pulmonary disease (COPD). This is a statistical review of the paper.
focusing primarily on the design, methods, analysis, reporting and interpretation of the results.

1. Abstract, Results:
   a. Report both point and interval estimates for the odds ratios in all comparisons.
   \textit{We added 95\% confidence intervals where they were missing.}

   b. Report all p-values to the same number of decimal places and apply this consistently.
   \textit{We used two decimals places throughout the manuscript.}

2. Methods:
   a. Provide the full search strategy as a supplemental file or appendix.
   \textit{As explained above, we based our search on existing systematic reviews and on complimentary PubMed related articles searches. We describe this search with more details in the method section and explained which databases were covered by these systematic reviews. In addition, we provide a list of systematic reviews that we used to identify relevant articles.}

   b. Who did the review of articles and what were their qualifications?
   \textit{In the revised paper, we describe in more detail who did the review and what their qualifications were. The revised text reads as:} “\textit{We retrieved all reports of potentially eligible trials. Two reviewers (research fellows with medical doctor degrees and 1 and 3 years of research experience, respectively) independently assessed them and determined their inclusion or exclusion. If the two reviewers disagreed even after discussion, a third reviewer (epidemiologist with MD and PhD degree) arbitrated.}” (…)\textit{ “}…"

   c. For the assessment of study quality, provide some rationale for the choice of the items used.
   \textit{We describe in more detail why we selected these criteria. The revised text reads as:} “\textit{We assessed if the method of randomization, concealment of random allocation and whether inclusion criteria were specified in order to judge whether confounding was controlled for by randomization and/or restriction. We recorded blinding of treatment providers and patients to judge the presence of information bias and we recorded}”
whether an intention-to-treat analysis was reported to assess if randomization was maintained throughout the analyses.”

d. How was the reliability of the data abstraction process assessed? Consider using Kappa statistics to measure agreement between reviewers.

*In the revised paper, we describe in more detail the data extraction process: “For each trial arm (2–4, depending on the trial), one reviewer extracted the number of patients with at least one exacerbation during follow-up and the number of patients with no exacerbation during follow-up (2×2, 3×2, or 4×2 tables). A second and, in case of disagreement, third reviewer, checked the data extraction for correctness.” We did not use Kappa statistics because the data extraction process was done by one reviewer and checked for correctness by the other reviewers.*

e. Provide the rationale for choosing logistic regression for analysis. This approach uses the fixed-effects model and also uses equal weight for all studies. Why not use a weighted approach?

*We chose a logistic regression model because of its transparency as compared with other methods for network meta-analysis. In the main analysis we used a fixed-effects model, indeed. However, studies are not weighted equally because we used quasi patient-level data. This means that, for example, a trial with 1000 patients contributed 1000 observations to the analysis whereas a trial with 50 patients contributed just 50 patients. Thus bigger trials contributed more to the analysis.*

3. Results: See comments 1.a. and 1.b.

*We corrected the paper accordingly.*

4. Discussion:

a. The Discussion needs to acknowledge some of the related recent reviews looking at safety issues of LABAs and SABAs:

We agree that outcomes other than exacerbations should be discussed. However, the debate about the safety of LABAs and SABAs relates mostly to asthma patients and it is unclear how this translates to COPD. We are, therefore, a bit hesitant to refer to these recent papers. Also, the outcome in these asthma studies was exacerbations, which is the outcome in our work anyway. But together with the next comment (4b) we believe that the discussion gains from a more thorough discussion of other outcomes. We, therefore, added a new paragraph to the discussion:

“Exacerbations is not the only outcome that should inform the decision for or against adding inhaled corticosteroids to long-acting bronchodilators. Other outcomes should be considered as well. A recent systematic review found no risk reduction in terms of mortality if an inhaled steroid was added to a long-acting bronchodilator.[11] Health-related quality of life was statistically significantly better after combined treatment (difference of -1.64 units on St. Georges Respiratory Questionnaire; 95% CI -2.28 to -1) but the effects were well under the threshold representing a minimal important difference (4 units).[11] Arguments for treating COPD solely with long-acting beta-agonists are the substantially lower costs and lower risk for adverse effects such as pneumonia, oral candidiasis, or loss of bone density compared to combination treatment[11, 14, 54, 61]. Integrating and presenting this complex information about benefits and downsides of inhaled drug treatments is challenging. One approach is to use decision aids that are particularly valuable for value-sensitive decisions where the balance of benefits and downsides is not straightforward. As the decision is, in the case of inhaled drug treatment for COPD, only between two treatment options (bronchodilator(s) versus combined treatment), information about the comparisons with placebo could now be excluded for simplicity since offering no treatment is not in the best interest of the patient. Whether patients benefit from such an informed decision making requires, however, testing in additional trials.”
b. The review should also balance the benefits against the harms. What is needed is a review that looks at both the relative benefits and harms in order to clarify the relative risk-benefit ratios of the drugs.

*We fully agree. Together with the safety issues we discuss this balance in more detail in the revised paper as stated just above.*