List of neuraminidase reviews with peer review comments and responses relevant to review A159

1. Neuraminidase inhibitors for treatment and prophylaxis of influenza in children: systematic review and meta-analysis of randomised controlled trials. Shun-Shin M, Thompson M, Heneghan C, Perera R, Harnden A, Mant D. BMJ. 2009 Aug 10;339:b3172. doi: 10.1136/bmj.b3172. Review.

   Documents:
   Appendix 1 Anonymised peer review comments.

2. Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis. Jefferson T, Jones M, Doshi P, Del Mar C. BMJ. 2009 Dec 8;339:b5106. doi: 10.1136/bmj.b5106. Review.

   Documents:
   Appendix 2 Anonymised peer review comments.

3. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ. Cochrane Database Syst Rev. 2012 Jan 18;1:CD008965. doi: 10.1002/14651858.CD008965.pub3. Review.

   Documents:
   Appendix 3A Anonymised peer review comments and author responses.
   Appendix 3B Anonymised feedback comments and author responses.

4. Neuraminidase inhibitors for preventing and treating influenza in healthy adults. Jefferson T, Jones M, Doshi P, Del Mar C, Dooley L, Foxlee R. Cochrane Database Syst Rev. 2010 Feb 17;(2):CD001265. doi: 10.1002/14651858.CD001265.pub3. Review known as A047.

   Documents:
   Appendix 4 Anonymised peer review comments and author responses.

5. NIHR Health Technology Assessment Programme Project: 10/80/01 - Update and amalgamation of two Cochrane Reviews: neuraminidase inhibitors for preventing and treating influenza in healthy adults and children

   Documents:
   Appendix 5A Anonymised peer review comments.
   Appendix 5B Response to peer reviewers.

6. Jefferson T, Jones MA, Doshi P, Del Mar CB, Hama R, Thompson MJ, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database of Systematic Reviews 2014, Issue 4. Art. No.: CD008965 DOI: 10.1002/14651858.CD008965.pub4.Review known as A159.

   Documents:
   Appendix 6A Anonymised peer review comments.
   Appendix 6B Response to peer reviewers.

7. Jefferson T, Jones M, Doshi P, Spencer E, Onakypoya I, Heneghan C. Oseltamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. BMJ.2014.017746.R3

   Documents:
   Appendix 7 Anonymised peer review comments and author responses.

8. Heneghan CJ, Onakpoya I, Thompson M, Spencer EA, Jones M, Jefferson T. Zanamivir for preventing and treating influenza in healthy adults and children: systematic review of clinical study reports. BMJ.2014.017753.R3

   Documents:
   Appendix 8 Anonymised peer review comments and author responses.

Appendix 1 Anonymised peer review comments

BMJ/2009/692921 [First Submission]

Neuraminidase inhibitors for the treatment and prophylaxis of influenza in children: a systematic review and meta-analysis of randomised controlled trials

Matthew Shun-Shin, Matthew James Thompson, Carl Heneghan, Anthony Harnden

Name: [redacted]

Position [redacted] Bond University Australia

This is an updated Cochrane systematic review of zanamivir and oseltamivir for influenza treatment and prophylaxis in children. It was first published in the Cochrane Library in 2003 and updated in 2007 with minor changes in personnel, but I could find no evidence it was published in print journals. It is clearly of great interest at the moment. The recent update is substantial – there are now more data from trials to increase the precision about some of the effects and harms of these drugs on children with seasonal influenza. The principle findings centre on the fact that effect sizes are very modest.

Neuraminidases are not dramatically effective at relieving symptoms. Influenza is clearly nasty – and a large minority of patients are suffering symptoms well into their 5th day. The drugs make a difference of somewhere between half to a day-and-a-half. The benefits of prophylaxis are also modest – with a NNT of 13 children treated to prevent one household contact spreading the flu. But oseltamivir is also unpleasant – the NNH (vomiting) is 20. This is rather reminiscent of the pros and cons of treating acute respiratory infections with antibiotics.

The paper is written very well, and is clear and direct. It is perhaps slightly long for the message, and could be rearranged with some stuff put on the Web for the cognoscenti. The main points are well brought out.

There is just one point that might be made more emphatically: the studies are on laboratory-confirmed influenza (and a lot of cases were Influenza A in most studies). In real life, more cases of influenza-like illness (ILI) cloud the issue. This MAY mean that the drugs are more effective in the current A/H1N1 epidemic (higher attack rate because of the high susceptibility of the population), (this point is well made), but MUCH LESS effective in 'seasonal' influenza, when influenza is confused with so much other viral illness.

It is probably the interpretation of these data that will raise discussion, with people divided into the Treat, and Don't, camps.

Name: [redacted]

Position: Consultant Senior Lecturer in Primary Health Care

Many thanks for asking me to review this interesting paper authored by an internationally respected group whom I know. They are to be congratulated on the timely submission of this highly topical article, though this does need to be squared with my first comment.

Major comments

1. I think the authors should be more clear in their introduction and the ‘what is already known’ section about the originality of the article in relation to the 2005 Cochrane Review.[1] To my mind, they need to state clearly why they felt a further review was necessary (e.g. to check that the Cochrane results were up to date?). The decision to publish in the BMJ may be based in part on whether this article is seen as a rapid update of the ongoing Cochrane Review or a new stand-alone piece of work. If publication is thought suitable, it would be ideal if the authors of the adult Cochrane review[2] were contacted to see if they too have conducted an update so that both papers are published together.

2. From the way the results are presented, I would guess that the trials present two very similar outcomes for symptom resolution (median days to resolution/alleviation of symptoms and median days to resolution of illness) given in Table 4. While this should be acknowledged, I think the current presentation is

confusing to the reader who may wonder why two such similar outcomes are given and whether they differ. It may be possible to prioritise one (with reasons) and only present these results. As a minor point related to this, the authors do not pool these data but do not state why.

3. Under strengths and weaknesses, why do the authors think that their strategy could have missed published trials? As a second point, I would distinguish the strengths and weaknesses of the review from those of the participating studies.

4. It would be particularly clinically helpful if some natural history data (from the placebo arms) of the trials could be given and, if possible, not just median illness duration since these data are likely to be positively skewed, but the time taken for 75% or 90% of children to recover.

Minor comments

1. To my mind there was an issue about whether the 'prevention of transmission' trials are clinically sufficiently homogeneous to pool results given the differences with which the index cases were treated (some given placebo, others same as neuraminidase under investigation, others pre-determined neuraminidase).

2. I think that the authors could usefully review their results section with an eye on their use of the term 'significant' in relation to statistical or clinical significance. Giving examples is hampered by the lack of manuscript page or line numbers but one specific example is under the paragraph 'effect of treatment on time to resolution of symptoms…' where in the last line they refer to a result that is unlikely to be a chance finding but appears relatively unimportant clinically.

3. Under effect of treatment on change in asthma severity, should the first study referenced by 16 (not 15)?

4. Under safety and tolerability, I think many clinicians would be interested to know more about the incidence of nausea, vomiting and diarrhoea. Can any further data be given here?

5. I think the authors should state why they decided not to do a funnel plot (may be because the number of studies was small).

6. Although already stated in Table 1, for convenience, I think Table 4 should also state which antiviral was under investigation and as a very minor point I would reverse the order of the antiviral and control columns.

(1) Matheson NJ, Harnden A, Perera R, Sheikh A, Symmonds AM. Neuraminidase inhibitors for preventing and treating influenza in children. Matheson Nicholas J, Harnden Anthony , Perera Rafael , Sheikh Aziz , Symmonds Abrahams Mkael Neuraminidase inhibitors for preventing and treating influenza in children Cochrane Database of Systematic Reviews : Reviews 2007 Issue 1 John Wiley & Sons , Lt 2007.

(2) Jefferson T, Demicheli V, Di PC, Jones M, Rivetti D. Neuraminidase inhibitors for preventing and treating influenza in healthy adults. Jefferson Tom , Demicheli Vittorio , Di Pietrantonj Carlo , Jones Mark , Rivetti Daniela Neuraminidase inhibitors for preventing and treating influenza in healthy adultsCochrane Database of Systematic Reviews : Reviews 2006 Issue 3 John Wiley & Sons, 2006.

Name: [redacted]

Position: Sr Researcher [redacted], science writer in several media.

General

I added in appendix a letter with my views on the use of NAI, which I never sent. While cost-effectiveness means that sick patients don't get expensive but effective treatment, large resources have been wasted to the illusions of containment of what was soon recognizable as a mild flu (I predicted the flu to be mild in a newspaper article that appeared 48 hours after the breaking news). I should be glad by this review, that adds to the evidence of poor effectiveness of NAI. However, this evidence is not new.

I feel this review adds too little to the existing reviews. I would suggest an update of the existing Cochrane review of treatment and chemoprofylaxis in children with a narrative report in the BMJ summarising the additional information of the recent studies. This would be short, but serve the same purpose of information

of the public. The present review is too little and too much. It adds too little for specialists, who should know the available evidence, it is too much for general medical interest.

Originality and importance

There are two reviews in the Cochrane database, one describing health effects in adult populations and one in children, referenced in refs 22 and 23. The Cochrane review of treatment in children has been executed by the same centre and supervised by the same supervising author (AH). This state of the art is essential, and should have been summarised in the introduction. It should have been made clear what the new information adds to the existing review. This is little. The added studies decreased the uncertainty of the effects of chemoprofylaxis among children (which are identical to similar studies in adults). The added treatment study in asthmatic children is important, but it is a single study.

The summary of "what this study adds" is oversold. It has been known that NAI reduces the duration of disease between 0.5 and 1.5 day in children and adults in winter flu. It will be less in the present mild pandemic flu.

Chemoprofylaxis is now a passed station (and has always been insane, see appendix). It was known from the existing reviews that the NNT of chemoprofylaxis to prevent one case of symptomatic flu was around 10 in adults and in children (a single study) (bulletpoint 1). The effect estimate in the population of children of asthmatics is important (bulletpoint 2), given that asthma is a risk factor for complications, but is based on a single study (ref 15). This finding could easily be covered by a short narrative report.

The effect of preventing complications in the seemingly more severe winter flu is known in adults (bulletpoint 3), reviewed in the Cochrane review of adults (ref 23). The NNT to prevent one hospitalization and one pneumonia in adults belonging to risk groups exceeded 100. To detect differences in the very rare serious complications in healthy children, the power needed is so large, that 'the value of added information' becomes an issue. With scarce research budgets, such a study would not be high on my priority list. The main aim of NAI, except for improving the quality of life of Roche and Glaxo shareholders, is to protect healthy people against the hypothetical complications of the hypothetical deadly flu looming somewhere behind the horizon, which will always be difficult to prove in evidence based medicine.

Methods, result and discussion

The actual paper builds on an existing high quality review and meta-analysis, adding one treatment study and three chemoprofylaxis study. There is a slight inconsistency: in the discussion, the studies are presented as poor, while in the results they are described as moderate (I suggest moderate is the correct description). I have no doubts that the present update is of identical high quality as the original study, executed by the same group. If the editors find that the message is still sufficiently important, given the present health scare, I would surely agree with publication. I would demand a honest representation of the state of the art in the introduction and what this new review adds.

Appendix

The true costs of saving hypothetical life years in pandemic health scares

For treatment strategies, rates of flu related morbidity and mortality in previously healthy adults are far too low to make treatment by NAI a cost effective option. It was soon clear that the actual swine flu was benign. Serious cases are the first to be detected. Younger people were infected first, but they travel more than the old and disabled. High child mortality by Mexican flu has to be interpreted against an under five mortality of 5% in Mexico (unacceptably high for this middle income country). The reasons for this high under five mortality are identical to what kills young flu patients: poorly nourished kids living in crowded and unsanitary conditions. Adding these up, all pointed to a mild flu from day one.

The rationale behind the large scale chemoprofylaxis by NAI (neuraminidase inhibitors) was containment of the flu epidemic, and is based on ring vaccination of smallpox. However, the typical symptoms of contagious smallpox are (or were) easily recognizable and dreaded by all. A single vaccination protected rapidly expanding social rings around the affected person and strangled the virus. Compare this to flu. The incubation period is shorter, flu is infectious before telltale symptoms develop, many flu cases remain mildly symptomatic or

asymptomatic. In the beginning of an epidemic, still rare flu cases are difficult to detect against a background of many infectious diseases causing similar symptoms. The rings of potential contacts is very small, NAI are only effective when administered and resistance develops rapidly. Containment seems only possible in the virtual world of the simulation model. If the present containment procedures slowed the epidemic, this slowing down will be counted in hours, not days. If this slowing down saved life years or health is unlikely.

The costs of the administrative flu circus, the serotyping and isolation procedures of healthy adults and the administration of prophylactic drugs for an essentially benign flu, are unknown to me. The resources, spent to containing an uncontainable but trivial flu, have been lost to strengthening primary child health care (in Mexico), improving care of the elderly, expanding respiratory care facilities, useful in potential true respiratory health crises. In the USA, 30000 people are yearly killed by guns; in 2004 (the most recent year available), 73 people were murdered by handguns in the UK, 11344 in the USA (http://www.ichv.org/Statistics.htm). Governor Schwarzenegger declared the state of emergency after two flu deaths. Health scares divert the attention from the true tragedies.

In health economy, money is saved life years. The pandemic of swine flu panic was then modestly lethal, not by the flu but by the wildly exaggerated public health response. Resources wasted to ineffective but high profile interventions profitable to industry and health bureaucracy were lost to more effective life saving interventions.

BMJ/2009/692921- Fast track provisional acceptance

Message: BMJ/2009/692921

Neuraminidase inhibitors for the treatment and prophylaxis of influenza in children: a systematic review and meta-analysis of randomised controlled trials

Dear Dr. Thompson

Thank you for sending us this paper and giving us the chance to consider your work, which we enjoyed reading. We are pleased to say that, with the rapid help of some excellent reviewers and the manuscript committee, we have decided to offer fast track publication in the BMJ, as long you are willing and able to revise the paper as we suggest in the report below.

We will need the revised manuscript back by **this Monday**, ie within 48 hours of your receiving our decision (plus this weekend). Once the final paper has been accepted and edited we will need the proofs returned within 24 hours.

We will aim to publish it within four weeks of its registration. Please upload the revised version as a Word document via your author area at

our online editorial office (http://submit.bmj.com) - do not resubmit the manuscript as a PDF because our system will not be able to process that.

All original research in the BMJ is published with open access. The full text online version of your article, if accepted after revision, will be the indexed citable version (full details are at http://resources.bmj.com/bmj/about-bmj/the-bmjs-publishing-model), while the print BMJ will carry an abridged version of your article soon afterwards.

While the revised and accepted paper is being edited we would like you to write an abridged version of the article for the print BMJ - what is essentially an evidence abstract called BMJ pico. For examples please see http://resources.bmj.com/bmj/authors/article-submission/bmj-pico-abridged-

research-articles. I will email you separately the appropriate template for abridging a systematic review in this way. Please be reassured that it doesn't take long to complete this.

As explained in this editorial (Groves, T, Godlee F. Innovations in publishing BMJ research. BMJ 2008 337:a3123 doi:10.1136/bmj.a3123), we're using BMJ pico to increase readership of research articles in the print BMJ and to give authors more control over the abridging. We have just completed surveys of authors and readers, showing that both groups like this format.

Meanwhile, I'm looking forward to seeing your revised article on Monday.

With best wishes and many thanks

Trish Groves

Report from the BMJ's manuscript meeting

We are able to accept only a small proportion even of the good articles submitted to us. A little over 10 % of articles reach this stage, and to do so they have to have passed preliminary screening by one or more of the editors, have received sufficiently positive external peer review, and have been discussed at the manuscript meeting.

At the manuscript meeting each article is discussed by the Editor or deputy, the rest of the BMJ's international team of research editors, and two invited advisers: one statistician and one clinical editorial adviser. As well as the scientific merits of the paper we take into account each paper's originality and interest to a general readership in comparison with other submitted papers. We take reviewers' reports fully into account too, but the final decision on acceptance or rejection of a paper rests with the Editor.

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Members of the committee were: Elizabeth Loder (chair), Doug Altman (senior statistics editor), Lucy Cheppell (editorial adviser), Trish Groves.

Decision: provisional acceptance (**fast track**).

Detailed comments from the committee:

* first and foremost, please revise your paper by responding to the comments by the reviewers. You will find these at our online editorial office (at http://submit.bmj.com) in your author area, under this manuscript number. We would like you to address all of the reviewers' comments, apart from the suggestion to contact Jefferson et al regarding an update of their Cochrane review of neuraminidase inhibitors in adults: although this might be ideal we didn't think it was strictly necessary now

* it wasn't nearly clear enough that this study is an update of your 2005 Cochrane review. This should be stated in the abstract, introduction (with the 2005 review being cited as one of the first references in the list), and methods, and the results should explain clearly what has changed or been confirmed since that 2005 review.

* small presentational points raised by Doug Altman: * the point of meta- analysis is the pooling, so it's not appropriate to dwell on the outcomes from individual primary studies (and particularly not in the abstract) * in table 1 the column heading "outcomes measured" should be changed to the more accurate term "outcomes reported".

Appendix 2. Anonymised peer review comments

BMJ/2009/726562 [Third Submission]
Neuraminidase inhibitors for preventing and treating influenza in healthy adults
Tom Jefferson, Mark Jones, Peter Doshi, and Chris B Del Mar

Comments
The revised article is in many places much improved and clearer. The authors have addressed most of my comments appropriately. There are still some important concerns that remain unaddressed, however. A few responses were a little abrupt and short (probably due to the amount of comments in total that had to deal with, and the short time-scale), and on a few occasions they have either not taken my advice and I still disagree, or said that the issue has been clarified in the text but I can't find it. These can be summarised as follows:

- It is still not clear how the hazard ratios have been calculated for each study in Figure 5. The extraction of hazard ratios is notoriously difficult for meta-analysis, and the authors simply report that they converted the medians of treatment groups into hazard ratios … but how, and what are 'medians of treatment groups'? They refer to a paper by Parmar et al., but this does not discuss how to convert median (survival time I presume) into a hazard ratio and a confidence interval. The authors need to be clear how they did this, as much hinges on the significant meta-analysis of hazard ratios reported. I would like to see their response to this please.
- At the start of the discussion the authors state that there is evidence of modest benefit for most illness – a reduction of the illness by about a day. But where does this conclusion come from? How is the 'one-day' extrapolated from the meta-analysis results – perhaps I have missed something? I can only see hazard ratios reported, but how are this risk ratio converted to a reduction in duration? Please be clear for the reader, as this is a crucial statement.
- I asked that the authors clearly show (in the paper, not just a web-extra) (i) how their new meta-analysis results differ from their previous meta-analysis results, and (ii) to show in a sensitivity analysis how the inclusion/exclusion of the 8 unpublished studies changes the effect on complications. I don't see either of this done in the new version, and I think the authors leave out the 8 studies entirely as they can't verify the age range of included patients. To help address this, as a reader I would prefer to see a summary table with a row for each meta-analysis performed and columns denoted the original 2006 result, the new meta-analysis result, and a final column explaining why there is a difference (if any) between the two meta-analyses (for example, exclusion of 8 studies that were unpublished with the reasons)

Also I identified two further issues that they should comment on:
- Is there any evidence of publication bias (small study effects) for any meta-analysis? Clearly unpublished studies are being excluded in some meta-analyses, so this is a natural question. Can the authors comment on this issue in the discussion please.
- Figure 4 shows considerable heterogeneity in all the meta-analyses. Thus, on page 7 when discussing these findings, the authors should clearly state that each pooled RR relates to the average RR across included studies. The RR in a single study may deviate substantially from this average value. So on average the RR is significant, but it may not be significant in a single study (due to the heterogeneity). See Higgins et al for further discussion on this issue.

Higgins JP, Thompson SG, Spiegelhalter DJ: A re-evaluation of random-effects meta-analysis. Journal of the Royal Statistical Society, Series A 2009, 172:137-159

Best wishes, [redacted] (SL in Medical Statistics)

- My comments on the revision:
Abstract: Section on search strategy does not make it clear they have not re-done entire search but rather have updated a previous search. Selection criteria: I am not sure this fully reflects the selection criteria. As I understand it, trials included not only healthy people exposed to influenza but also those with symptomatic influenza.

Results: The second paragraph really should provide results with and without those trials, especially since it's uncertain whether there is anything wrong with these data at all. The statement about "possibly underreported" seems out of place in the results section. They do not present evidence in the paper that supports this assertion.
Discussion: Use of the word "modest" is confusing, since it means different things to different people. Also, the focus on influenza symptoms in healthy adults does not reflect the findings on PEP. The statement "should not be regarded as mandatory" does not reflect the findings of this study.

"What this study adds": the first point was shown in a previous review so they should say "confirms findings of our previous review" and they should also mention benefits on PEP. I do not know what they mean about "insufficient evidence for or against" adverse events. There is clear evidence of nausea and no evidence of other major adverse events.

Introduction: Unclear whether blocking neuraminidse blocks viral entry of all sorts of viruses or just influenza viruses. Is there any reason to think it might work better on a particular type of influenza virus? Do "influenza-like" viruses express neuraminidase? Casual readers won't know. Third paragraph: change "this criticism centered on one study" to "centered on one paper" -- since that paper incorporated data from several studies.

Methods: First 2 paragraphs should explain why they did not do a full and complete search this time around but simply did an updated search. I am not sure why they excluded studies of experimental influenza (appendix says they did but why? we should be told in this section a bit more about the inclusion and exclusion criteria).

Maybe I'm missing it, but I can't find inclusion and exclusion criteria listed anywhere for either the efficacy/effectiveness searches or the AE searches. They are not in Box 1. Rather sly that they now call their exclusion of the studies in the Kaiser paper a "sensitivity analysis". If this paper has to stand alone, they need to make clear what the findings of their previous review were and that one of the included papers (which summarized 10 studies) was criticized and therefore they have redone the review. I still do not know what they mean by "raw data".

Results: End of the first paragraph doesn't make sense. Might mention "not able to unconditionally provide the data as quickly as we needed it to update this review". It's not clear why they were "obliged" to exclude the Kaiser data. It is my understanding that a number of the other studies contained a mixture of healthy adults and those with comorbidities, so I wonder why they are applying stricter crtieria to these studies. They say they excluded it "after sensitivity analyses" so why don't they present those analyses, eg the "before" as well as the "after"?

Treatment: Could they provide us with the information on the magnitude of change in the duration of ILI?

Quality of the evidence: the information in this section doesn't seem to deal with the "quality" of the evidence, eg how accurate it is. Rather, it seems to deal with the "completeness" of the evidence. They still have not explained clearly who has to report to these different AE databases, what period of time reports cover, how reports are compiled, and so forth. Perhaps a table would help, with rows being the 3 AE databases and columns representing their various attributes and requirements. I don't understand what they mean when they say "irreconcilable differences". They mention a "nonsignificant trend for causing diarrhea" for oseltamavir, but that's not a statement we ordinarily allow. The confidence interval for prophylaxis of ILI in healthy adults is very wide and crosses 1. Rather than saying they found "no effect" I think it would be more accurate to say there is "insufficient evidence to support or refute the effect of NIs on prophylaxis of influenza-like illness."

Discussion: Should discuss benefits on all examined outcomes. Currently they are selective. Again I think use of the term "modest" is a value judgment that is in the eye of the beholder. They could simply say by what degree it reduces symptoms and say that in their opinion this is a modest benefit at the level of the individual. Their review did not examine possible public health benefits or harms of the drug so I think the parts of the discussion that deal with the role of NIs in pandemic influenza are outside the scope of this work.

Also in discussion: rather than say "inaccessible to proper scrutiny" it might be more even-handed to say" we were offered access under conditions we thought were unacceptable" and point readers to the accompanying Doshi article.

Somewhere they should acknowledge that Hayashi's comments on the original review raise questions about the conduct of the orginal review and Cochrane reviews in general. As I said in my comments about the abstract, they don't really present evidence that harms are underreported so I do not think this should be a conclusion. I do not understand what they mean in the second paragrahp on page 11...do they mean they register the AEs before 3 years are up and then stop?

The summary findings do not flow from the review. The review was focused on reporting on trial evidence on a number of outcomes: prophylaxis of ILI (findings were of no effect); prophylaxis against laboratory confirmed influenza (findings were of an effect); post-exposure prophylaxis in exposed adults (results reported in the text but not in a Forest plot); alleviation of influenza symptoms (finding of an effect); effect on complications of influenza requiring antibiotics (even without the disputed studies the visual effect is quite striking...a fairer statement would be insufficient evidence to confirm benefit). The summary should stick tightly to findings on these outcomes.

Box 1: It's unclear why they excluded studies at the level of record screening and eligibility.

[redacted] review: The revised article is in many places much improved and clearer. The authors have addressed most of my comments appropriately. There are still some important concerns that remain unaddressed, however. A few responses were a little abrupt and short (probably due to the amount of comments in total that had to deal with, and the short time-scale), and on a few occasions they have either not taken my advice and I still disagree, or said that the issue has been clarified in the text but I can't find it. These can be summarised as follows:

- It is still not clear how the hazard ratios have been calculated for each study in Figure 5. The extraction of hazard ratios is notoriously difficult for meta-analysis, and the authors simply report that they converted the medians of treatment groups into hazard ratios … but how, and what are 'medians of treatment groups'? They refer to a paper by Parmar et al., but this does not discuss how to convert median (survival time I presume) into a hazard ratio and a confidence interval. The authors need to be clear how they did this, as much hinges on the significant meta-analysis of hazard ratios reported. I would like to see their response to this please

- At the start of the discussion the authors state that there is evidence of modest benefit for most illness – a reduction of the illness by about a day. But where does this conclusion come from? How is the 'one-day' extrapolated from the meta-analysis results – perhaps I have missed something? I can only see hazard ratios reported, but how are this risk ratio converted to a reduction in duration? Please be clear for the reader, as this is a crucial statement.

- I asked that the authors clearly show (in the paper, not just a web-extra) (i) how their new meta-analysis results differ from their previous meta-analysis results, and (ii) to show in a sensitivity analysis how the inclusion/exclusion of the 8 unpublished studies changes the effect on complications. I don't see either of this done in the new version, and I think the authors leave out the 8 studies entirely as they can't verify the age range of included patients. To help address this, as a reader I would prefer to see a summary table with a row for each meta-analysis performed and columns denoted the original 2006 result, the new meta-analysis result, and a final column explaining why there is a difference (if any) between the two meta-analyses (for example, exclusion of 8 studies that were unpublished with the reasons)

Also I identified two further issues that they should comment on:
- Is there any evidence of publication bias (small study effects) for any meta-analysis? Clearly unpublished studies are being excluded in some meta-analyses, so this is a natural question. Can the authors comment on this issue in the discussion please.
- Figure 4 shows considerable heterogeneity in all the meta-analyses. Thus, on page 7 when discussing these findings, the authors should clearly state that each pooled RR relates to the average RR across included

studies. The RR in a single study may deviate substantially from this average value. So on average the RR is significant, but it may not be significant in a single study (due to the heterogeneity). See Higgins et al for further discussion on this issue.

Higgins JP, Thompson SG, Spiegelhalter DJ: A re-evaluation of random-effects meta-analysis. Journal of the Royal Statistical Society, Series A 2009, 172:137-159Robin Ferner did a review of first resubmission

[redacted] West Midlands Centre for Adverse Drug Reactions

The authors revise their views of neuraminidase inhibitors after a further systematic review of evidence, and contact with a pharmaceutical company.

1. If I have understood the paper and supplementary material, the authors faced three difficulties.

2. First, in an initial review, results for 8/10 efficacy trials were taken from a secondary report by Kaiser and others, who worked for Roche. When challenged on this by Hayashi, the current authors properly tried to obtain the necessary data from Roche, but were offered it on condition that they signed a confidentiality agreement, one of whose clauses was that they should keep the existence of the confidentiality agreement confidential.

3. Secondly, the efficacy data were derived from cohorts in which a high proportion of influenza-like illness (ILI) was demonstrated to be caused by influenza virus. Although effectiveness (measured by improvements for those with ILI) was not demonstrated, efficacy (measued by improvements in outcomes for those exposed to or suffering

4. Thirdly, the data on harms are difficult to find, the numbers of spontaneous reports and their provenance are unclear, and deductions from the FDA data are distorted by loss of reports from elsewhere in the world.

5. These are extremely important issues, and will be of great topical interest to readers of BMJ.

6. The authors show that adverse reaction reporting rates fell as usage of neuraminidase inhibitors increased. In fact, these data reflect the 'Weber effect,' whereby reporting rates for newly introduced drugs are much higher than for established drugs. [see, for example, http://www.lareb.nl/documents/PWS2003_1088.pdf]

7. The data on harms are not entirely clear. This is partly because the authors do not explain their criteria for accepting or rejecting papers on adverse reactions [page 6, para 3; chart 2]. Inclusion of FDA data from spontaneous reports suggests that the authors were looking for 'signals' as well as confirmatory studies from randomized trials or case-control data. They omit, however, related data from UK Medicines and Healthcare products Regulatory Agency.[UK Suspected Adverse Drug Reaction (ADR) Analysis Influenza antivirals - oseltamivir (Tamiflu) and zanamivir (Relenza) 29 October 2009] A table, or possibly a Venn diagram, could help.

8. For these and other reasons the data on harms are confusing.(See point 4 above). The suggestion that 'Governments should set up studies to monitor the safety of oseltamivir' [page 8 para 4] is not helpful. Some governments have already set up spontaneous reporting schemes. They, and other non-randomized studies, will be subject to bias. And randomized trials for safety are difficult, because they have to be very large to detect rare adverse effects.

9. Two important questions are unanswered, namely: (1) do neuraminidase inhibitors reduce the incidence of serious complications from influenza (esp pandemic influenza), and (2) do they improve outcome in patients who are already seriously ill.

10. The authors advocate the use of neuraminidase inhibitors in life-threatening illness, but they should perhaps demand evidence of efficacy here (since a clinical trial showing benefit would be easiest in this group). There is a feeling among intensivists and BMJ editorialists [BMJ 2009;339:b2698]that they do work. But neither group is infallible.

11. There are one or two quirks. NPAEs are referred to on page 7, but defined on page 8, for example.

[redacted]

Appendix 3A Anonymised peer review comments and author responses.

Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children

Oct 2010

reviewer and criticism / suggestion	Our response
contact editor CE	
1 it is hard to read the protocol and understand all the methodologies proposed for the review	We are not sure what the problem is here. We have attempted to write this as clearly as possible.
2 Perform two reviews ONE - more clinical, ready for immediate use and friendly to readers ANOTHER addressing the methodological issues (bias tables , impacts of including unpublished studies on the results, comparisons with the previous reviews , methodological recommendations for further Cochrane reviews etc	 We agree. We have limited this review to the ONE There remains the need for some focussed methodological work. We propose to establish a Cochrane Working group to address non-published pharmaceutical trials. This will be proposed at Keystone Colloquium in Oct 2010.
3 agrees with (?all) reviewers' comments:	
also extra comments	
- the protocol :- a short description of the local and systemic symptoms of influenza in the background (for someone who does not make the difference between common cold and influenza) is missing	added
- mortality seems to me a PRIMARY outcome during an influenza epidemic	too rare to have as a primary outcome.
- consider the primary outcome "symptomatic relief" for anyone who seems to have influenza without lab or only for subjects with laboratory confirmed influenza for the treatment studies?	we make this distinction as 'true (ie lab-confirmed) influenza' or 'Influenza-like illness, ILI'
- in the secondary outcomes how would be "interruption of transmission" defined?	any outcome that showed that treating one population decreased the incidence (compared to control)
- "unit of analysis issues" include RR, AR, RD, NNT - these should be moved to "measures of treatment effect"	not sue what is meant here...
PR 1 [redacted]	
1 Too many hypotheses in Table 1 and 2 – so many as to invoke Bonferroni corrections?	These are the methodological ones. We propose removing these to make the review easier to conceptualise, and move these methodological issues to a Methods Working Groups (proposed).
2 minor stuff	
- make it clearer that this was a merging of two previous reviews.	Done
. - Outcome measures: inconsistency between the stated primary outcomes (1. Symptom relief ...) and the statement that: "We plan to focus on complications and adverse events". Perhaps the latter could be reworded as: "We will pay particular attention to complications and adverse events, including "compliharms".... "	Done
- Statistical Methods state that: "When no heterogeneity is detected, we will perform a random-effects meta-analysis" Of course, if there is no heterogeneity, fixed and random effects give the same results, but it is odd to state it this way.	Done – reworded to clarify
PR2 [redacted]	
1. The title does not reflect the goal of the study. Although the conclusions of the study are not available which may ultimately alter the title, the title does not elude to why the study was done, why it is relevant.	The title is written in standard form: Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. We don't agree that this should change (nor should it contain the 'results'
2. In the "Description of the Condition" the description of influenza can be modified to read as: "Occasionally patients with influenza will develop complications including pneumonia, otitis media and dehydration that may be due to effects of the influenza virus itself and/or to associated secondary bacterial infections."	This sentence has been clarified.
3. In the section "How the Intervention might work" It is well understood how NI's work on the virus itself. ts efficiency on the population would also be appropriate to discuss here. In this section describe how the intervention works in decreasing symptoms and shedding in an infected individual and the reducing the spread of the virus among a population (cases within nursing homes, etc).	we have added this to the bottom of the paragraph: "Any treatment that reduces the excretion of virus from infected people might be useful public health measure to contain an epidemic. Indeed this is the basis of using neuraminidase inhibitors during the threatened H1N1 outbreak of 2009.
4. Papers from groups in Japan have looked at the neurological effects and actions of oseltamivir. This could also be included in the importance.	Yes this planned. Indeed one of us (Rokuro Hama) wrote some of those papers. Of course we will get the information from the trials of the pharmaceutical industry. So this does not need to be stated explicitly.
CR1 [redacted]	
1 suggestions for providing more detail in the Introduction.	Done see 1st paragraph of "Why it is important to do this review "
CR2 [redacted]	
it is very important to be matter of fact in this protocol, keeping to the guidelines for Cochrane reviews multiple detailed suggestions	 some changes to the text to achieve some of these suggestions.

Appendix 3B Anonymised feedback comments and author responses

From: [redacted]>
Date: 15 December 2010 18:51
Subject: Neuraminidase inhibitors for influenza - HTA project
To: "cdelmar@bond.edu.au" <cdelmar@bond.edu.au>, "jefferson.tom@gmail.com" <jefferson.tom@gmail.com>, Carl Heneghan <carl.heneghan@dphpc.ox.ac.uk>

Hi
I picked up Carl's Twitter request for comments on your draft protocol "Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of unpublished data". So, here are my two comments on the content.

The title confused me: I expected it to be a review of unpublished trials to complement your review of published trials. It would be longer but clearer if you could call it "Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of clinical study reports for published and unpublished trials".

The section "How the intervention might work" could be reorganized along the lines of:
0) Metabolism: oseltamivir phosphate (OP), Tamiflu, is the pro-drug of oseltamivir carboxylate (OC), the effective form. OP dissociates in the gastrointestinal tract to form oseltamivir (OT) which is absorbed and metabolised into OC by hepatic carboxylesterase (h-CE).
1) Reducing the ability of the virus to penetrate the mucus in the very early stage of infection (Bhatia 2007; Matrosovich 2004; Moscona 2005; Ohuchi 2006).
2) Inhibiting neuraminidase, which enables influenza viruses to exit host cells (Liu 1995; Moscona 2005).
3) Central depression by OT (Hama 2008) may cause hypothermia (Ono 2008).
4) Inhibition by NIs of human sialidase may cause abnormal behaviour (Li 2007).

You have obviously put a huge amount of work and expertise into developing the protocol, and have an even bigger task ahead to complete the review. Congratulations for taking this on.

Best wishes
[redacted]

Reply

Thanks for the constructive comments.

1. We have re-titled the Protocol to address this concern (and that of feedback from GSK, see below);
2. We have re-examined the "How the intervention might work" section, but made only small adjustments in the interest of keeping this section short;
3. We are not sure what problems you might have had printing the pdf file, and hope they are resolved with this new version.

Contributors

Chris Del Mar

From [redacted] 24 February 2011

Summary

From: [redacted]>
Date: 24 February 2011 12:48
Subject: oseltamivir
To: jefferson.tom@gmail.com

I've read your Intervention Protocol: Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of unpublished data. And may be you can be interested in this letter I wrote to de BMJ: http://www.bmj.com/content/340/bmj.c789.extract/reply
1. Early use of oseltamivir does not reduce swine flu mortality, Juan C. Vergara, MD. Intensive Care Unit, Hospital Cruces. 48901 Barakaldo. Spain

As you say, in July the National Pandemic Flu Service started providing oseltamivir to anybody who telephoned with a plausible set of symptoms. From 23rd July to 1st December, the National Pandemic Flu Service (NPFS) in the UK, has provided more than one million courses of antiviral medication. By that time the Spanish Health Secretary General, José Martínez Olmos, at the Congress of Deputies, announced that only 6.000 patients (most of them hospitalised) had received oseltamivir in Spain. At the end of January there have been 411 deaths reported due to pandemic (H1N1) 2009 in the UK, and about 300 in Spain. That means 6.7 and 6.5 deaths per million, respectively. These data create serious doubts about the real utility of early use of oseltamivir in preventing deaths from Influenza A H1N1.
http://www.nhsdirect.nhs.uk/article.aspx?name=SbSwineflu
http://www.congreso.es/public_oficiales/L9/CONG/DS/CO/CO_411.PDF

Competing interests: None declared

Yours sincerely;
[redacted]

Reply

Thank you for your interest.

Contributors

Chris Del Mar

From [redacted], GSK, UK, 30 March 2011

Summary

GSK comments on Cochrane Collaboration protocol: neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of unpublished data

General:

- The term ‘unpublished data’ is used extensively in the protocol. However, it does not appear to be clearly defined either in the protocol or in Jefferson’s comment in the 15 Jan 2011 edition of the BMJ. Additionally, the term ‘unpublished data’ is misleading. It appears the Cochrane Group use this term interchangeably with Clinical Study Reports, regardless of whether a primary manuscript is available for a given study. We suggest this is clarified or preferably replaced, especially since the term appears extensively in the protocol including the title. Readers are likely to use the terms ‘unpublished data’ and ‘unpublished trials’ (trials for which no primary publication appears in the scientific press) interchangeably. A suggested replacement is ‘Clinical Study Reports’ since this term is not easily misinterpreted and is clearly defined in Jefferson’s BMJ comment.

- The ‘scope of clinical trial data’ are defined in Jefferson’s BMJ 15 Jan 2011 comment, as mentioned above (i.e. definitions for clinical study reports, raw data, unpublished trial, published trial, regulatory data). It would seem important that these and any other definitions introduced in the protocol are included in the protocol.

Description of Intervention

- This section incorrectly describes Relenza as ‘nebulized zanamivir’. Relenza is formulated in Rotadisks containing foil blisters with a powder mixture of zanamivir and lactose. Relenza is administered by oral inhalation using a breath-activated device called the Diskhaler. Earlier clinical studies explored several methods of administration, including nebulized and intranasal routes, but marketing approval in nearly all countries is currently available only for oral inhalation via Rotadisk/Diskhaler.

Types of Studies

- To meet the objective of providing a comprehensive review of neuraminidase inhibitors in preventing and treating influenza, it would seem appropriate that clinical trials from all sources (including sponsors other than industry) be included in this meta-analysis. Please clarify if this is your intent.

Outcome Measures

More details should be provided on the outcome measures section in the final protocol.

- For example, broad outcome measures are stated in the protocol, but specific endpoints are not provided. The primary and secondary endpoints of the meta-analysis should be clearly defined in the final protocol.

    -
    - e.g.1. A stated primary outcome in the treatment studies is ‘symptom relief’. Does this refer to ‘the time to alleviation of symptoms’ or ‘reduction in symptom score’ or another endpoint? Time to alleviation of clinically significant symptoms was the primary endpoint used in the majority of GSK treatment studies.

    -
    - e.g.2. Another stated primary outcome is ‘Harms’. Please provide the specific endpoints. Will this refer to ‘incidence of most common AEs’ or ‘incidence of common SAEs’, ‘incidence of complications’ or another endpoint? It is not clear if ‘harms’ are the same as ‘compliharms’. It is not clear what specific events will comprise compliharms.

- Prophylaxis studies: Several types of prophylaxis studies were conducted by GSK: household prophylaxis (post-exposure prophylaxis), community prophylaxis and outbreak control in nursing homes, and as such the designs and/or endpoints are different. It is possible to measure ‘prevention of onset of influenza in contacts’ in these studies, but not ‘reduction in viral spread from index cases’ in the majority of prophylaxis studies.

- Hospitalizations: As studies were generally conducted in the setting of acute uncomplicated influenza, limited hospitalisation data were collected, and are available only for some studies.

- Extracting compliharms: There is a statement that 'AEs are reported for all participants while complications are only reported for infected subjects'. This statement is not accurate for GSK trials. AEs are reported for all study participants. However, AEs of ILI were not collected in the treatment studies unless the symptoms were considered to be worse than expected for the normal progression of illness. Without knowing the specific safety endpoints, it is unclear whether this will affect the outcome of some of the harms analyses.

Data collection and analysis:

- The protocol indicates that clinical study reports will be requested (minus participant identification). In fact many documents for each study will need to be redacted not just to remove participant identification, but any personally identifiable information including author and investigator identification.

- Missing Data. The protocol states "At the participant level (i.e. within a trial) we will not make any assumptions about missing data." This is not possible, because an analysis of data that is collected in a trial can only be done in the context of assumptions about potential mechanisms that led to data being missing (e.g., missing completely at random, or missing at random).

- Meta-analysis Method. Little detail is given in the protocol. The protocol states that "Whether or not heterogeneity is detected, we will perform a random effects meta-analysis. Random-effects methods will be used to compare the dichotomised outcomes (RR and absolute risk reduction (ARR) for efficacy and safety)." There are several different Random Effects methods available (Bayesian or frequentist, DerSimonian & Laird or Maximum-likelihood or REML), and different approaches to handling rare events (various "corrections" to include trials with zero counts). Furthermore, would random-effects methods also used to compare the continuous outcomes?

- Fixed-effects Model. The protocol also states that fixed-effects models will be used in a sensitivity analysis. No details are given with regard to which fixed-effects models will be used. There are several fixed-effects models available including Inverse Variance, Mantel-Haenszel, and Peto's method. The appropriate method used should also depend on the outcome measures (dichotomous vs. continuous; relative vs. absolute). The approach and choice of models for sparse data and rare events should be provided. Furthermore, various methods in the framework of fixed-effects model may be explored to evaluate the robustness of the results.

- Hazard Ratio. The protocol states "We will convert medians of treatment groups into (log) hazard ratios (estimating the variance of these) to enable meta-analysis of time to event outcomes." Although hazard ratio (HR) is a standard analysis and widely recommended approach for time-to-event data in clinical trials, the HR analysis may not be suitable for the Relenza studies with relatively short follow-up time because the assumption of proportional hazards required for the proportional hazards model may not hold. GSK did not follow this approach for the original analysis due to the concern stated above. Further the clinical and regulatory interest centred on differences in the time to alleviation not in the relative hazard between treatments. The above issues would be best addressed by using subject level rather than summary data, which GSK have offered to provide to the Cochrane Group.

- Analysis Populations. The protocol does not specify which populations will be used for the various analyses, for example, intent-to-treat or influenza-positive or other. We believe that influenza positive population is appropriate, especially for the efficacy analysis using time to alleviation of influenza symptom as a primary endpoint consistent with the prescribing information for Relenza.

- Study Duration. No details are given in the protocol with regard to how studies with different follow-up times will be handled.

- Trials with no Events. No details are given in the protocol with regard to how to deal with trials in which there are no events (such as death). By excluding studies with no events will make the event appear more common than it actually is. There are various techniques: Bayesian approach, continuity correction, combining similar trials to avoid having any components of the analysis that have no events.

- Sensitivity Analyses. Sensitivity analyses using different outcome measures, statistical models and/or continuity correction factors to assess the robustness of the results are strongly encouraged.

Reply

General:

- 'unpublished data'. We agree that this term is confusing, and are attracted to the proposal of using 'clinical study reports' instead.
- We have attempted to ensure all terms are clear.

Description of Intervention

- Description of zanamivir (Relenza): we have corrected 'nebulized zanamivir' to 'powder inhalation'.

Types of Studies

- Yes, we intend to comprehensively review clinical trials from all sources (including sponsors other than industry). This intent is clear from the subsection "Electronic Searching " under the "Search methods for identification of studies" Section.

Outcome Measures

- Our specified outcomes are those of interest to patients, and their clinicians and policy-makers. They are therefore likely to be broader than the more specific endpoints selected by trialists. The purpose of Cochrane Reviews are usually to set clinically relevant review questions, and search the literature (or other sources) for answers to them. Sometimes answers to some questions are not available, and this is also documented. Where possible we report outcomes as pre-specified in the trial protocols, or as pre-specified in the review protocol, or otherwise reported as a post-hoc analysis.

    - e.g.1. 'symptom relief' may refer to 'the time to alleviation of symptoms' or 'reduction in symptom score', or any other endpoint (including 'area under the curve of symptom score and time').

    - e.g.2. 'Harms' include common adverse events (AEs) as well as serious AEs. We agree about the confusion of harms and complications, and have tried to capture the totality of these with the neologism 'compliharms' to avoid classification errors between their different labellings.

- Prophylaxis studies: We understand that it is possible to measure 'prevention of onset of influenza in contacts' in some GSK studies, but not 'reduction in viral spread from index cases' in others.

- Hospitalisations: We understand that hospitalisation data may only be available for some studies. However patient hospitalisation is usually classified as a serious adverse event therefore we expect to identify hospitalisations (not reported separately) in that way.

- Extracting compliharms: Your statement that 'AEs of ILI were not collected in the treatment studies unless the symptoms were considered to be worse than expected for the normal progression of illness' underlies the complexity of analysing AEs and complications (our 'compliharms'). We have noted in the protocol that the limitation of complications only reported for the infected patients is relevant to the Roche trials only.

Data collection and analysis:

- We are interested that not only subject identification would be required to be removed from any documents of clinical study reports, but also information personally identifying authors and investigators. We wonder why.

- Missing Data. We have removed this statement.

- Meta-analysis Method. DerSimonian & Laird method will be used. Note that in the case of zero cells (e.g. no events in one group) the RevMan software (which we will use for the analysis) automatically adds 0.5 to each cell of the 2×2 table for any such study. There are no continuous outcomes specified in this review.

- Fixed-effects Model. Mantel-Haenszel method will be used except in the case of sparse data, in which case Peto's method will be used (as recommended in the Cochrane handbook).

- Hazard Ratio. We note the concerns with this outcome hence we will also consider analysis of this outcome as a continuous outcome noting that the data is likely to be skewed. We will use the inverse-variance random-effects method for this analysis.

- Analysis Populations. All analysis will be using the intent-to-treat population as this is the most methodologically rigorous and clinically relevant.

- Study Duration. We have specified in the protocol, where appropriate, that we will report outcomes for the on-treatment and off treatment time periods. If data is not available in the clinical study reports for any time period of the study then we will write to the relevant manufacturer to request the missing data.

- Trials with no Events. As stated above the RevMan software automatically adds 0.5 to each cell of the 2×2 table for any such study.

- Sensitivity Analyses. We note this point and agree. Where appropriate, a realistic sensitivity analyses will be conducted.

Contributors

Chris Del Mar

Feedback from [redacted], 30 January 2012

Summary

Dear Tom Jefferson,

I read your review about NI for prevention and treating influenza with interest. It's an important work. In the chapter "Why it is important to do this review" I found a small mistake concerning the worldwide stockpiling of oseltamivir which is mentioned to be "CHF 7.6 billion worth of oseltamivir (JACK 2009)". This would be an enormous amount "prior (!) to the emergence of influenza A/H1N1 in 2009". But Andrew JACK wrote in the cited Financial Times (May 13, 2009): "Governments around the world had stockpiled 220m treatments to date, swelling sales since the start of 2003 to SFr7.6bn, largely on the basis of preparation for a pandemic virus that has yet to appear." So 7.6 billion SFr represent sales and not stockpiling.

Wolfgang Becker-Brueser (physician and pharmacist)

From [redacted], 2 February 2012

Summary

I am writing to comment on the recently updated meta-analysis by Jefferson and colleagues published through the Cochrane Collaboration and to request clarifications on several points, as well as to suggest some additional analyses that would be helpful in terms of taking greater advantage of this useful database. While I fully support access of Jefferson and other interested investigators to all of the published and unpublished data from the RCTs of oseltamivir and zanamivir for further analyses, this analysis only focuses on RCTs in ambulatory patients with uncomplicated influenza (the vast majority of whom were previously healthy) and on the period before the 2009 H1N1 pandemic. Consequently, I would urge these investigators to extend their efforts to other populations and datasets examining the risks and benefits of using neuraminidase inhibitors (NAIs) for treatment and prophylaxis. Furthermore, the authors should acknowledge the limitations of their analyses more explicitly and avoid inappropriate extrapolation to populations and influenza events that the RCTs did not adequately address. Differences in disease pathogenesis related to virus and host factors, as well as time to treatment, have important effects on the utility of antiviral agent interventions. My specific comments and recommendations for additional analyses follow:

1. Use of Intention to Treat (ITT) and ITTI-Infected Groups. The exclusive focus in the current treatment analysis on the ITT population is a readily rectified shortcoming. Outcomes in all three groups of relevance (ITT, ITT-infected, and ITT-noninfected) should be presented, so that readers can examine both clinical effectiveness and efficacy for the key endpoints, as well as events in those without documented influenza. Because NAI treatment would not be expected to provide any benefit in non-influenza illness, not presenting the ITT-infected outcomes in the analysis underestimates possible beneficial drug effects. Assessment of the non-infected group provides a valuable control and also enables a determination of whether there was a potential drug-disease adverse interaction of NAI treatment in non-influenza patients. Of note, our earlier pooled analysis of physician-diagnosed lower respiratory tract complications leading to antibiotic use found a significant benefit of oseltamivir in the influenza-infected patients but not in those enrolled in whom influenza infection was not detected by culture or serology [Kaiser 2003].

2. Sample size considerations. Severe outcomes of influenza infection are sufficiently uncommon in previously healthy people that even large RCTs or combining multiple RCTs would be very unlikely to detect them with confidence. The same point applies to very uncommon endpoints like microbiologically documented bacterial complications and rare adverse effects of treatment. Consequently, conclusions that

there is no evidence (from trials) that NAIs reduce the risk of pneumonia, hospitalizations, deaths are overstated, as the evidence considered in this analysis is insufficient to properly address these questions.

The US CDC has estimated age-related influenza-related hospitalisation and mortality rates for both seasonal epidemics and the 2009 pandemic [Shrestha 2011]. Jefferson and colleagues should use such event estimates and others to make calculations of the necessary sample sizes to detect reductions in these severe outcomes with NAI therapy in a controlled RCT across a range of clinically relevant effect sizes (e.g., 20%, 35%, 50% reductions). In a related fashion, they should also provide more quantitative estimates for their ability to detect such outcomes with their existing database and comment more precisely on their power to capture particular endpoints.

3. Complications in ambulatory patients. Other clinically relevant endpoints in these previously healthy and at-risk persons warrant investigation. With regard to influenza-related complications, the most frequent in previously healthy children and adults are respiratory tract infections (otitis media, bronchitis) leading to antimicrobic use. These are usually not severe and typically not microbiologically documented with respect to etiologies, but physician-diagnosed complications leading to antibiotic use is an outcome that has important clinical and public health implications (i.e., cost, antibiotic resistance, side effects) and also is sufficiently frequent to demonstrate effects of antivirals. We showed such a benefit in adults in our earlier pooled analyses of the then available RCT data on inhaled zanamivir [Kaiser 2000] and oral oseltamivir [Kaiser 2003]. The oseltamivir effect was confirmed in a recent meta-analysis [Hernan 2011], and another recent Cochrane report confirms an effect on otitis media in children [Wang 2011].

Given the large amount of data available to the investigators, it would be a valuable contribution to also explore the clinical outcomes in greater detail and to clarify the use of terms like severe outcomes. Although uncommon in the populations enrolled in these RCTs, endpoints such as radiographically documented pneumonia, microbiologically documented infections, and hospitalisation or death are clear and should be listed separately in those with or without proven influenza infection. Because of the importance of hospitalizations as an endpoint, it would be helpful to examine not only all-cause hospitalizations but also relevant subgroups based on likely causation (e.g., events in which influenza was documented or likely implicated including exacerbations of co-morbidities vs others like accidents, elective surgeries, conditions unlikely to be influenza-related). In addition to these events, exacerbations of underlying conditions (e.g., asthma, COPD, diabetes, CHF) are of medical importance in influenza outpatients with co-morbidities and should be examined.

4. Data from observational studies. Typically the patients who are most at risk of severe outcomes (older people, infants and young children, those with underlying chronic conditions) are not included in RCTs. In this regard, the current analysis is limited to placebo- or active-controlled RCTs largely done in previously healthy persons and does not consider the multiple observational studies from different countries that have consistently showed protective effects against severe outcomes like pneumonia and hospitalisation, particularly in those with co-morbidities, as well as reduced mortality if patients have been hospitalised. A considerable amount of new treatment data was generated in many countries during the 2009 H1N1 pandemic that found timely NAI treatment to be associated with a lower risk for intensive care admission and death (reference list available upon request).

While such data and analyses are weaker than RCT data and subject to bias, these observational studies address key endpoints in at-risk and seriously ill populations, including patients admitted to a hospital at the time of initiating therapy, that the available RCTs cannot and do not address. Furthermore, the standard of care has evolved such that placebo-controlled RCT in such patient groups would not be acceptable to investigators or ethics committees. The decision by Jefferson and colleagues not to consider and critically analyze the large amount of observational data with modern techniques means that they are not incorporating key information and many important patient groups in which the available data suggests medically important benefits from early NAI therapy. Such findings from observational data can inform antiviral treatment in more severely ill patients when no other data are available. As discussed above, not to include observational data means that conclusions of no effect on uncommon events or no severe adverse events being detected are almost inevitable. This should be made explicit in the design and the conclusion of the current report.

4. Influenza diagnosis and serologic results. The Jefferson report raises questions about the possible inhibitory effects of oseltamivir therapy on influenza-specific serologic rises and introduction of bias into the outcomes analysis. Further analyses might help to assess these possibilities. They should compare the primary endpoint of illness alleviation between the oseltamivir and placebo subgroups that were culture-positive (irrespective of serologic findings) at enrolment, and separately those that were culture-negative but had serologic evidence of infection.

Of note, one prior study of oseltamivir treatment in pandemic 2009 H1N1 patients, although not in seasonal influenza patients, suggested that early treatment could reduce antibody responses [Cowling 2010]. Jefferson and colleagues should examine the age-related frequencies of HAI seroconversions and the GMT titer rises in those with influenza-culture positive illness and separately in those with such HAI rises in absence of culture positivity. Of course, if still available, it would be interesting to test the culture-negative enrolment samples by RT-PCR.

The RCT data were generated over multiple seasons in which different influenza A and B viruses were circulating. Influenza B neuraminidases are generally less susceptible to oseltamivir carboxylate and several observational studies indicate that oseltamivir is less effective in influenza B- than influenza A-infected children [Sugaya 2007; Sato 2008]. It would be useful to examine the primary outcome in relation to virus type (A vs. B) and if possible A subtype (H3 vs. H1) in those with documented infections to expand on this point.

5. Other treatment endpoints of interest. Since those enrolled in the RCTs were outpatients, it would be useful to explore other endpoints that reflect patient recovery and impacts on the healthcare system (e.g., nonscheduled return visits for complications or adverse events). Perhaps more important than the time to alleviation endpoint used in the registrational trials might be the times to resumption of usual activities and return to pre-morbid status.

The authors raise the possibility that oseltamivir might have non-specific antipyretic effects, and one animal model study has also suggested possible adverse immunomodulatory effects of oseltamivir in RSV infection [Moore 2007]. Consequently, it would be interesting to examine the course of fever resolution (a much earlier event than cough resolution) and of symptoms in oseltamivir- and placebo-treated patients with and without documented influenza infections. In addition, it would be valuable to examine the correspondence (or lack thereof) between influenza virologic measures (e.g., enrolment virus titer, time to culture negativity, change in viral titers over time) and symptom resolution measures in both oseltamivir and placebo groups.

Various cost-effectiveness analyses on NAI therapy in low-risk populations have been published with widely divergent outcomes, largely depending on the input assumptions. Using this large database, a more refined analysis that incorporates both the direct and indirect (productivity losses) costs of influenza would be informative.

6. Adverse events with treatment. With regard to drug tolerability, it is important to examine not only the frequencies of reported adverse events but also assess indicators of their severity and interference with compliance (e.g., symptom days, patient reported severity, premature cessation of study drug).

Comparisons of AEs in the placebo groups across zanamivir and oseltamivir studies need to be interpreted with caution, since these studies were performed in different influenza seasons viruses and locations, with different protocols and case record forms, and by different investigators. Only one head-head RCT of treatment comparing these drugs has been published to date to my knowledge but the design did not include placebo only groups [Duval 2010]. In particular, comparisons in children (page 24) need to be age-adjusted as there were major differences in those enrolled into the zanamivir (5 years and older) and oseltamivir trials (1 year and older), and the frequencies of gastrointestinal manifestations are much higher in younger children with influenza and other acute illnesses.

7. Prophylaxis endpoints of interest. The analysis of prophylaxis outcomes and the associated discussion requires clarification. The statement on page 5 says: “The FDA has also not allowed an indication for interference of viral transmission within households (the key concept behind post-exposure prophylaxis).” The

key concept behind post-exposure prophylaxis is prevention of illness in exposed persons, and the primary endpoint in most prophylaxis studies has been symptomatic, laboratory-confirmed influenza illness. FDA and other regulatory agencies have approved both NAIs for post-exposure prophylaxis in households and also for longer duration pre-exposure chemoprophylaxis [reviewed in Khazemi 2009].

The Jefferson analysis seems to focus exclusively on the effect of chemoprophylaxis in "preventing the spread" of influenza, with endpoints presumably determined by evidence of culture or serologically confirmed infection irrespective of illness. While this is one endpoint of interest in such studies, the primary outcome of medical interest is prevention of influenza illness in those exposed. There is abundant RCT data, as well as observational data from the 2009 pandemic, that both inhaled zanamivir and oral oseltamivir have both statistically significant and medically important effects on preventing influenza-specific illness. Of note, the development of serologic evidence of infection without illness is advantageous in those receiving chemoprophylaxis, as it likely is an immunizing event that protects against future infection and illness by that strain. In addition several oseltamivir RCTs have shown significant but lesser effects on influenza infection in prophylaxis recipients [Welliver 2001; Hayden 1999]. The authors should present all of the relevant endpoints in their analysis of the prophylaxis trials.

8. Adverse effects with prophylaxis. The prophylaxis studies are particularly useful in assessing drug tolerability as symptoms of acute illness present in treatment studies are not confounders and there is a more prolonged duration of drug exposure. However, it is essential to examine not only the frequencies of reported adverse events but also indicators of their severity and possible interference with compliance (e.g., symptom days, patient reported severity, premature cessation of study drug).

For example, the Jefferson posting states that "Similarly, a published prophylaxis trial (Hayden 1999a, known by its trial ID WV15673/WV15697) describes headache as having "occurred in similar proportions of subjects in the three groups (39 to 47 per cent)." but indicates that Japanese regulatory documents reached a different conclusion. My own review of the adverse event tabulations from our 6-weeks prophylaxis study (tables provided by the sponsor) indicates that the proportions of subjects reporting headache (not otherwise specified) that might have been related to study drug (unrelated reports excluded) during the treatment phase were similar across the placebo (N=116, 22.4%), oseltamivir 75 mg once (N=124, 23.8%), and oseltamivir 75 mg twice (N=132, 25.4%) daily dose groups [Hayden 1999]. Most of these reports indicated mild or moderate intensity and were self-limited. As indicated in the published paper [Hayden 1999], study withdrawals for AEs or illness occurred infrequently across these same groups (N=10, 1.9%; N= 8, 1.5%; N= 7, 1.3%). Of note, the specified causes for AE-related withdrawals included three reports of headache associated with other symptoms in the placebo group. In contrast, there were no reports of headache as reason for the withdrawals receiving oseltamivir; gastrointestinal complaints accounted for withdrawals in 4 of 8 oseltamivir 75 mg and 3 of 7 oseltamivir 75 mg twice daily recipients. The total numbers of patients with premature study withdrawal for any reason was 21 (4.0%), 17 (3.3%), and 16 (3.1%) across the three groups, respectively. Overall, severe AEs were reported in 82 (15.8%) of placebo, 75 (14.4%) of oseltamivir 75 mg, and 77 (14.8%) of oseltamivir 75 mg twice daily recipients. We were unable to include these details in the paper because of space limitations, but my interpretation remains that no excess of clinically relevant oseltamivir-related headache occurred during this study. This type of detailed AE analysis incorporating severity measures provides necessary context in interpreting the possible importance of AEs.

9. Peer review. The questions raised and opinions expressed in this and earlier Cochrane reports on NAIs by Jefferson and colleagues have resulted in debate and sometimes confusion among practitioners and policy makers regarding the appropriate use of NAIs in seasonal and pandemic influenza responses. Given the importance of these issues, it would be helpful for any future updates to have proper independent review before posting or publication by the Collaboration, as the Cochrane methodology of publication and then independent peer review is not well understood by many people.

Thank you for the opportunity to provide comments. I look forward to seeing the responses from Dr. Jefferson and his colleagues on these points.

Sincerely,

[redacted]

Reference List

1. Cowling BJ, Chan KH, Fang VJ, Lau LL, So HC, Fung RO, Ma ES, Kwong AS, Chan CW, Tsui WW, Ngai HY, Chu DW, Lee PW, Chiu MC, Leung GM, Peiris JS. Comparative epidemiology of pandemic and seasonal influenza A in households. N Engl J Med 2010 June 10;362(23):2175-84.

2. Duval X, van der WS, Blanchon T, Mosnier A, Bouscambert-Duchamp M, Tibi A, Enouf V, Charlois-Ou C, Vincent C, Andreoletti L, Tubach F, Lina B, Mentre F, Leport C. Efficacy of oseltamivir-zanamivir combination compared to each monotherapy for seasonal influenza: a randomised placebo-controlled trial. PLoS Med 2010:7(11):e100362.

3. Hayden FG, Atmar RL, Schilling M, Johnson C, Poretz D, Paar D, Huson L, Ward P, Mills RG. Use of the selective oral neuraminidase inhibitor oseltamivir to prevent influenza. N Engl J Med 1999 October 28;341(18):1336-43.

4. Hernan MA, Lipsitch M. Oseltamivir and risk of lower respiratory tract complications in patients with flu symptoms: a meta-analysis of eleven randomised clinical trials. Clin Infect Dis 2011 August 1;53(3):277-9.

5. Kaiser L, Keene ON, Hammond JM, Elliott M, Hayden FG. Impact of zanamivir on antibiotic use for respiratory events following acute influenza in adolescents and adults. Arch Intern Med 2000 November 27;160(21):3234-40.

6. Kaiser L, Wat C, Mills T, Mahoney P, Ward P, Hayden F. Impact of oseltamivir treatment on influenza-related lower respiratory tract complications and hospitalizations. Arch Intern Med 2003 July 28;163(14):1667-72.

7. Khazeni N, Bravata DM, Holty JE, Uyeki TM, Stave CD, Gould MK. Systematic review: safety and efficacy of extended-duration antiviral chemoprophylaxis against pandemic and seasonal influenza. Ann Intern Med 2009 October 6;151(7):464-73.

8. Moore ML, Chi MH, Zhou W, Goleniewska K, O'Neal JF, Higginbotham JN, Peebles RS, Jr. Cutting Edge: Oseltamivir decreases T cell GM1 expression and inhibits clearance of respiratory syncytial virus: potential role of endogenous sialidase in antiviral immunity. J Immunol 2007 March 1;178(5):2651-4.

9. Sato M, Saito R, Sato I, Tanabe N, Shobugawa Y, Sasaki A, Li D, Suzuki Y, Sato M, Sakai T, Oguma T, Tsukada H, Gejyo F, Suzuki H. Effectiveness of oseltamivir treatment among children with influenza A or B virus infections during four successive winters in Niigata City, Japan. Tohoku J Exp Med 2008 February;214(2):113-20.

10. Shrestha SS, Swerdlow DL, Borse RH, Prabhu VS, Finelli L, Atkins CY, Owusu-Edusei K, Bell B, Mead PS, Biggerstaff M, Brammer L, Davidson H, Jernigan D, Jhung MA, Kamimoto LA, Merlin TL, Nowell M, Redd SC, Reed C, Schuchat A, Meltzer MI. Estimating the burden of 2009 pandemic influenza A (H1N1) in the United States (April 2009-April 2010) Clin Infect Dis 2011 January 1;52 Suppl 1:S75-S82.

11. Sugaya N, Mitamura K, Yamazaki M, Tamura D, Ichikawa M, Kimura K, Kawakami C, Kiso M, Ito M, Hatakeyama S, Kawaoka Y. Lower clinical effectiveness of oseltamivir against influenza B contrasted with influenza A infection in children. Clin Infect Dis 2007 January 15;44(2):197-202.

12. Wang K, Shun-Shin M, Gill P, Perera R, Harnden A. Neuraminidase inhibitors for preventing and treating influenza in children. Cochrane Database Syst Rev 2012;1:CD002744.

Welliver R, Monto AS, Carewicz O, Schatteman E, Hassman M, Hedrick J, Jackson HC, Huson L, Ward P, Oxford JS. Effectiveness of oseltamivir in preventing influenza in household contacts: a randomised controlled trial. JAMA 2001 February 4;285(6):748-54

Submitter has modified conflict of interest statement: Disclosures to BMJ (Updated 4 June 2012)

[redacted] received lecture and/or consulting honoraria from GSK until 2002 and from Roche until 2005. Gilead Sciences from 1996-1999 and Roche from 1999-2005 provided grant support to the University of Virginia for oseltamivir studies on which he was PI. Similarly GSK provided grant support to the University of Virginia for zanamivir studies from 1994-2001. [redacted] served as medical officer in the [redacted] from 2006-2008 with funding provided to the University of [redacted]. Since 2008 to present the University of [redacted] has received funding from the [redacted] for his part-time work as influenza research coordinator at the Trust and through [redacted] for his work as consultant the [redacted]. From 2008-11 the University also received honoraria for his participation in the Neuraminidase Inhibitor Susceptibility Network which received funding from Roche and GSK. Since 2008 to present [redacted] has been an unpaid consultant to multiple companies engaged in the development or marketing of influenza antivirals including Roche and GSK.

[redacted] reports receiving compensation as a member of the scientific advisory boards of Novartis and Immune Targeting Systems, and has performed consulting work for Pfizer. Within the last 3 years, his group has been funded to perform laboratory assays or conduct clinical trials for Sanofi, GlaxoSmithKline, Protein Sciences Corp, Wyeth, PaxVax, Ligocyte, and Vaxinnate.

[redacted] reports no financial disclosures.

[redacted]

Reply

Response to [redacted] comments of 2 February 2012.

We thank [redacted] for his detailed feedback. However nothing he writes allays our basic concerns that:

(1) Despite the 16,000 pages we analysed, we currently only have access to a very limited dataset hence cannot carry out many of the analyses [redacted] suggests;
(2) Analyzing the "influenza infected" population in Roche oseltamivir trials, as [redacted] proposes, will lead to misleading results because the treatment groups are not comparable for this population;
(3) The observational studies [redacted] urges us to consider are generally of poor quality and only represent the small proportion of patients who are hospitalised with influenza;
(4) the Kaiser et al (2003) analysis is seriously flawed;
(5) Data have been selectively reported.

Below, we provide point-by-point responses to [redacted] concerns. (Please note that point 4 appears twice, to follow the numbering in [redacted] letter.)

1. Use of Intention to Treat (ITT) and ITTI-Infected [sic] Groups.

We agree, in principle, to conduct analysis using the ITT-Infected (ITTI) sub-population provided that it is appropriately selected by the results of testing completed before the start of the trial (for example by using only the results of viral culture or rapid testing before randomisation).

However we argue that this is not possible in Roche oseltamivir trials. In these trials, the selection of "infected" or "non-infected" was dependent on the results of serology that is affected by "use" and "non-use" of oseltamivir. And the selection of those with "serology-positive results" appears to have given advantage to the oseltamivir group. Hence the method of selecting the ITT-Infected population in the trials has fundamental flaws and therefore the results are less reliable than those obtained using the ITT population.

2. Sample size considerations.

The Kaiser et al analysis has a number of fundamental problems. First, analyses were performed on the ITT-Infected sub-population which we have shown to be non-comparable between treatment groups. Second, the authors analysed an outcome that was different to that pre-specified in the trials. In the trials, complications included otitis media and sinusitis but in the Kaiser et al paper these were not included. This is an example of selective reporting or "cherry picking". Third, complications were not objectively or consistently measured in the trials. Fourth, outcomes such as pneumonia and bronchitis could be either reported as a complication or as an adverse event according to a classification criteria we do not understand and is not discussed in the Kaiser et al paper. And finally the data from the 10 trials was not meta-analysed, rather, it was combined as if generated from one single trial.

We could potentially address most of these limitations (except for the third) but we have not been given access to the data despite repeated requests to the manufacturer. However we were able to compare hospitalisations as those data were available to us for the ITT population.

We found no evidence of effect on hospitalisations based on seven studies with a median placebo group event rate of 0.84% (range 0% to 11%): odds ratio (OR) 0.95; 95% CI 0.57 to 1.61, P = 0.86). This result is quite different to that reported by Kaiser et al based on the (non-comparable) ITT Infected population.

In terms of power analysis, to detect a significant difference at this level of difference of 0.84% (placebo) vs 0.80 % (oseltamivir), with alpha of 0.05 and power of 0.8, a RCT with approximately 800,000 participants is required.

3. Complications in ambulatory patients.

As we have illustrated above the Kaiser et al (2003) analysis has fundamental flaws that we cannot address because the manufacturer refuses to provide us with the data necessary to conduct a proper analysis.

Analysis of the "population with proven influenza infection" (ITT-Infected population) is not appropriate (see above). Data for the analysis of "population without proven influenza infection" are not available to us.

As we have shown above, the power to detect a difference in all-cause hospitalisation is very small hence to do a subgroup analysis on this outcome seems unwarranted.

The pharmacological/toxicological adverse effects of oseltamivir can be classified into two major types [3]. One is sudden type occurring during the hypercytokinemic state in the early phase of infection including sudden death [3,4], accidental death after abnormal behaviours and vomiting induced by the central depressing action of unchanged oseltamivir [4]. The second are delayed type of reactions including recurrence or exacerbation of influenza and/or other infection, diabetes, bleeding, renal impairment and delayed type neuropsychiatric reactions related to inhibition of the host's neuraminidase [3]. Sudden type adverse effects should be collected and analysed only during the early phase of influenza (for example, vomiting was only significantly increased within one day of treatment in the paediatric RCTs). However, delayed type adverse effects should be collected and analysed for a longer period to detect those reactions after a full course of treatment (for example the increase of pneumonia in the off-treatment period in the paediatric RCTs).

A recently published proportional mortality study has indicated that oseltamivir increases sudden type of death (odds ratio: 5.9) compared with zanamivir users by analysing all death cases among approximately 20 million 2009A/H1N1 influenza patients in Japan. This effect was also true for the comparison of oseltamivir users with non-users of antivirals [4].

4. Data from observational studies.

Observational studies during the 2009 H1N1 influenza outbreak have assessed the effects of oseltamivir on a selected population of hospitalised patients. These represent a very small proportion of the total population who get influenza. While subgroup analyses are important, it is important to not lose sight of the fact that the use and governmental stockpiling of oseltamivir is for its routine use in asymptomatic and symptomatic members of the community. Our review thus considers the evidence base that applies to the vast majority of people.

In addition, the studies [redacted] appears to be referring to are retrospective observational studies in which apparent treatment effects may be the result of an effective treatment but could also be due to confounding effects. Unfortunately there is no way to determine which of these possibilities is true. That is why drug regulators require evidence from RCTs to determine whether or not a drug is approved for use. According to the analysis by Jones and Hama [5], apparent protective effects against severe outcomes like pneumonia, hospitalisation and mortality are possibly derived from survivor treatment selection bias (or immortal time-bias). This is not an issue for randomised controlled trials because follow up begins at the time of randomisation which is the same for patients allocated to active drug and patients allocated to placebo. However in the case of observational studies treatment can begin at varying times (up to several days) after the onset of symptoms. Therefore a naive comparison that compares a binary outcome, such as death (or other adverse event), or time to an event (survival time) is at high risk of survivor treatment selection bias (also referred to as immortal time bias or simply time dependent bias). This bias can occur, for example, because patients who die early are not given the opportunity to receive treatment. In addition patients who are extremely sick may not be given the opportunity to receive antivirals because other treatments and procedures take priority. This bias can be addressed with an appropriate analysis however this has not been done in any of the observational studies of antiviral use for influenza that we have seen.

4. Influenza diagnosis and serologic results.

We do not have access to the data required to conduct all these analyses.

5. Other treatment endpoints of interest.

We do not have access to the data required to conduct these analyses (time to resumption of usual activities and return to pre-morbid status) using the ITT population.

By mentioning the evidence and possible mechanism of action for oseltamivir, we are arguing that fever alleviation and symptom reduction may not be caused by the reduction of viral load but may be the result of inhibition of host's immune functions including induction of cytokines and antibody production by inhibition of the host's neuraminidase in addition to central depression by oseltamivir.

Analysis of the population with documented influenza infection (ITT-Infected population) is not valid (see above). Hence we are unable to conduct a valid analysis in the influenza positive population and data for the influenza negative population has not been provided.

Antibody titre is one of the ways of selecting only subjects infected with influenza. However we have shown that the production of antibodies was consistently lower in the oseltamivir group compared to the placebo group in the treatment trials. Therefore the use of antibody production to confirm influenza in prophylaxis trials is not valid. Moreover comparison of the proportion with confirmed infection between the oseltamivir group(s) and the placebo group will provide misleading results.

Nor are “virus titre”, “time to culture negativity” or “change in viral titres over time” a true measure of viral load, because oseltamivir as a neuraminidase inhibitor may conceal positivity by inhibiting the influenza virus from leaving the surface of host respiratory cells (which are covered by a mucous layer on the surface of the cells).

6. Adverse events with treatment.

In principle we agree. However, there are many data that show the classification of severity is questionable: for example, we believe that psychosis or hallucinations should be classified as “severe” but this has not always been followed. Therefore, we are planning to propose using new classification methods for the analysis of adverse events in the next update of our review.

We agree that comparisons of adverse events in the placebo groups across zanamivir and oseltamivir studies need to be interpreted with caution.

We agree that the spectrum and severity of adverse events/reactions are different among age groups. Therefore, we propose analysing adverse events/reactions stratified by age, if possible, according to the data in the Clinical Study Reports or individual patients’ data in the next step of our systematic review.

7. Prophylaxis endpoints of interest.

As described on page 7 of our systematic review, the primary outcome measures for prophylaxis studies are:

1. Influenza (both symptomatic and asymptomatic and laboratory-confirmed) and influenza-like illness (ILI);
2. Hospitalisation and complications;
3. Interruption of transmission (in its two components, reduction of viral spread from index cases and prevention of onset of influenza in contacts);
4. Harms.

We did not meta-analyse data from the prophylaxis trials in this systematic review because the substantial documents for prophylaxis trials were obtained after the time-lock of April 12th 2011.

Due to the problems we have illustrated above on using virus titer to confirm influenza infection we plan to amend the primary endpoint for prophylaxis trials to influenza-like illness (ILI).

There is some fear that those with serologic negative infection without symptoms may be more easily infected with influenza virus in the future, because evidence from animal experiments shows that IgA antibody in the respiratory mucosa is reduced (to about 20% of the control group), while reduction of those of systemic IgG antibody (HI antibody) was slight and not statistically significant [6].

8. Adverse effects with prophylaxis.

We agree that the prophylaxis studies are particularly useful in assessing drug tolerability.

As we discussed above (“7. Adverse events with treatment”), there are many data that show the classification of severity is questionable. For example, we believe that psychosis or hallucinations should be classified as “severe” but this has not always been followed. Therefore, we are planning to propose using new classification methods for the analysis of adverse events in the next step of the review.

We mentioned the statement “occurred in similar proportions of subjects in the three groups (39 to 47 per cent)” as an example of reporting bias present in the paper ([redacted]’s reference no. 3; known by its trial ID WV15673/WV15697).

The numbers for headache are 47% (242/520) in high dose oseltamivir group, 43% (335/520) in low dose oseltamivir group and 39% (202/519) in placebo group. These proportions are not similar and show a significant linear trend of increase with oseltamivir dose (P = 0.013).

In addition, we would be grateful if [redacted] were to supply the definition of "drug related headache among headaches reported as adverse events"? In particular, how was it decided whether a headache was drug-related or not? We cannot suggest signs or symptoms to distinguish oseltamivir-induced headache from placebo-induced headache.

We propose analysing adverse events in clinical study reports, including those for prophylaxis trials.

9. Peer review.

We agree that there is confusion among policy-makers and practitioners but believe this to be justified: the data published and accessible to them appear to have some flaws that need to be resolved. We are encouraged by Dr [redacted]'s support for our obtaining all the data necessary to clear the confusion.

Cochrane systematic reviews are stringently peer-reviewed. Not only are they peer-reviewed by independent experts prior to publication, but the protocols are also peer-reviewed before being undertaken, to reduce a priori biases. In addition, protocols are available for comment from outside the internal review process – Dr [redacted] himself, or employees of Roche the manufacturer of oseltamivir, could have provided input about suggested alterations to the protocol which we would have been glad to receive. To this extent the peer-review process is more stringent than that employed by most other scientific journals.

RH, MJ, TJ, CDM, PD

References

[1] Moore ML, Chi MH, Zhou W, Goleniewska K, O'Neal JF, Higginbotham JN, Peebles RS Jr. Cutting Edge: Oseltamivir decreases T cell GM1 expression and inhibits clearance of respiratory syncytial virus: potential role of endogenous sialidase in antiviral immunity. J Immunol. 2007 Mar 1;178(5):2651-4.

[2] Hama R, Jones M et al. An overview of neuraminidase inhibitors on inhibitory effect of immune response including cytokines and antibody production (under submission for publication)

[3] Hama R, Fatal neuropsychiatric adverse reactions to oseltamivir: case series and overview of causal relationships. The International Journal of Risk & Safety in Medicine 20 (2008) 5-36. Available at http://npojip.org/english/no11.html

[4] Hama R, Jones M, Okushima H, Kitao M, Noda N, Hayashi K, Sakaguchi K. Oseltamivir and early deterioration leading to death: a proportional mortality study for 2009A/H1N1 influenza. Int J Risk Saf Med. 2011;23(4):201-15. Available at http://iospress.metapress.com/content/5257410g24403m68/fulltext.pdf

[5] Jones M, and Hama R et al. Survivor treatment selection bias in a cohort of 2009A/H1N1 influenza patients from Japan (manuscript in preparation)

[6] Takahashi E, Kataoka K, Fujii K, Chida J, Mizuno D, Fukui M, Hiro-O Ito, Fujihashi K, Kido H. Attenuation of inducible respiratory immune responses by oseltamivir treatment in mice infected with influenza A virus. Microbes Infect. 2010 Sep;12(10):778-83.

Contributors

Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ

Additional feedback from [redacted], 10 August 2012

Summary

I am writing to respond to the comments and questions raised by Jefferson and his colleagues to my letter of 2 February 2012 about their report published through the Cochrane Collaboration. While the authors have provided helpful clarifications to many points, I remain concerned about their selective approach to data analysis and presentation. Resolution of these issues is important in anticipation of future analyses by Jefferson and colleagues or by others. Many of their responses indicate that analysis of the cohorts with proven influenza infection (ITT-infected) are not appropriate, but further analyses of patient level data should be able to address their concerns (see below). Also they identify biases that could make oseltamivir look better but not those that could make it look worse than its effectiveness and tolerability likely are in reality. An impartial analysis would identify biases in both directions and attempt to deal with them in a balanced appraisal.

My specific comments and recommendations for additional analyses follow:

1. Use of Intention to Treat (ITT) and ITTI-Infected Groups. One obvious means of addressing the concern about selection bias in defining the ITT-infected (ITTI) population for analysis is to focus on those who were influenza virus-positive (irrespective of serologic results) at enrolment. These individuals (ITTI-virus) represented approximately 70-85% of those enrolled into the ITTI cohorts across the various RCTs.

In addition, those who were included in the ITTI group solely on the basis of seroconversion could be analysed separately to assess overall comparability in terms of symptom resolution and complications to those who were both virus-positive (ITTI-virus) and showed serologic rises. This might also help determine whether inclusion of data from virus-negative seroconverters would affect overall findings.

In contrast to the Cochrane statement that “And selection of those with “serology-positive results” appears to have given the advantage to the oseltamivir group”, it might alternatively be disadvantageous (bias toward the null) or neutral in effect. If oseltamivir is most beneficial in preventing lower respiratory tract (LRT) complications leading to antibiotic use in those in whom it also prevents seroconversion, as one might expect if its overall treatment effect varies between patients based on timing of administration, individual pharmacokinetics or other factors, then its protective effect on complications will be underestimated because the benefits in those for whom it prevents seroconversion will not be counted. If, on the other hand, treatment works effectively only in those infected who seroconvert and has little or no effect in those in whom it prevents seroconversion, this would increase the apparent benefit. However, the only way in which this sequence seems possible would be if late treatment does not interfere with seroconversion but early treatment does AND late treatment is more effective than early. This is biologically implausible and inconsistent with the observed effects on time to treatment for other outcomes, in which early treatment is associated with greater effects. Alternatively, if oseltamivir treatment has a similar effect on LRT complications in infected who seroconvert and those who do not, this would reduce the numbers in the treated group with and without outcomes in a non-differential way.

In addition to a possible non-specific immunomodulatory effect of oseltamivir on serologic responses or possible confounding effect of prior inactivated influenza vaccine which might blunt antibody responses in those with proven influenza (1), one explanation for the apparently lower seroconversion rate in oseltamivir recipients would be that some oseltamivir recipients had low viral replication levels at enrolment that were quickly reduced by treatment and did not stimulate antibody rises, so that in these persons treatment prevented seroconversion. If one assumes that clinical outcomes are linked to viral replication levels as other

reports suggest, such individuals would probably have shorter illness duration and also be less likely to develop LRT complications. Consequently, not counting them in the oseltamivir group would bias towards the null and under-estimate the effect of treatment on both illness resolution and complications. In this regard, comparing outcomes in the ITTI-virus seroconverters vs non-seroconverters would be of interest if sufficient numbers are available. Also, as stated previously, analysis of the serologic responses based on time from symptom onset to enrolment, including both frequency of seroconversion and observed titers rises in the ITTI-virus group compared to placebo, might help address this possibility.

If I have interpreted their report correctly, the post-hoc analyses by Jefferson and colleagues found an absolute difference of 3.4% in overall infection rates between placebo (68.9%) and oseltamivir (65.5%) groups across the studies they analysed (Figure 5, Table 17). This difference presumably approximates the fraction of virus-negative, non-seroconverting but possibly influenza-infected subjects in oseltamivir group. To what extent this difference might bias outcomes is uncertain, but its relatively modest size suggests that misclassification would not be a major confounder in either the ITTI or ITT-non-infected groups. Optimally in future studies more sensitive nucleic acid amplification testing will be used to detect infection by influenza and other respiratory viruses and facilitate more clear delineation of the groups of interest.

In summary, further analyses of the RTCs on oseltamivir and zanamivir, the outcomes in all groups of relevance (ITT, ITTI, ITTI-virus, and ITT-non-infected) are important and should be presented as fully as possible. As stated previously, separate assessment of the ITT-non-infected group provides a valuable control and also enables a determination of whether there was a potential drug-disease interaction of NAI treatment in non-influenza patients. As specific antiviral treatment would not be expected to provide benefit on illness resolution or complications in non-influenza illness, examining the ITT-non-infected groups allows this point to be tested directly. An analysis of 11 oseltamivir RCTs (2) confirmed lack of treatment effect on LRT complications in non-influenza-infected subjects compared to placebo. The failure to present outcomes in the ITT-infected or ITT-virus cohort underestimates possible beneficial drug effects, whereas full data presentation would enable readers to examine the event rates and magnitude of treatment effect sizes for key outcomes across all relevant groups for themselves.

2. Sample size considerations. The endpoint used in our pooled analysis of oseltamivir RCTs (3) was prospectively defined before the analysis was undertaken and was based on findings in our earlier study of zanamivir treatment effects (4) that indicated inhaled zanamivir reduced LRT illnesses leading to antibiotic prescriptions (RR, 0.60; 95% CI, 0.42-0.85), but not upper respiratory tract ones (RR 0.90; 95% CI, 0.63-1.27). The oseltamivir analysis used all studies available to us at the time, including unpublished clinical study reports, in order to avoid selection bias. The other endpoints of upper respiratory tract complications leading to antibiotic use (6.8% oseltamivir vs 5.9% placebo) and overall antibiotic use (14.0% oseltamivir vs 19.1% placebo; P <.001) were described in our 2003 paper (page 1760). Of note, the reductions in overall antibiotic use in influenza outpatients were similar for zanamivir (28%) and oseltamivir (27%) treatment. The limitations of the clinical diagnoses and retrospective approach used in these studies were described more fully in the earlier zanamivir paper (4). However, the simple pooled analysis we undertook in the oseltamivir paper did not correct for the higher proportion of influenza-infected, at-risk individuals in the placebo group, and this was a shortcoming. In any case, we pointed out this difference in the paper (page 1669) and presented the data by each group of interest (previously healthy or at risk) in Tables 3 and 4.

More importantly, our finding that early oseltamivir treatment reduced the likelihood of physician-diagnosed LRT complications leading to antibiotic use has been confirmed and extended (37% reduction in oseltamivir group; risk ratio 0.63 [95% CI, 0.48, 0.82]) in a subsequent meta-analysis (that controlled for pre-enrolment risk status and included events from the time of enrolment) of the same 10 RCTs included in our paper and one additional one (2). Furthermore, this analysis found that the unpublished trials for which Jefferson and colleagues apparently do not have data were found to be no more favourable to oseltamivir than the published ones. When only the two published trials in previously healthy persons were considered, the reduction in the 24-day risk of LRT complications treated with antibiotics was 65% (risk ratio, 0.35; 95% CI, 0.15, 0.82) in the oseltamivir arms.

3. Complications in ambulatory patients. Their comments on possible oseltamivir adverse events, including sudden death and neuropsychiatric adverse events (NPAEs), raises important points about the effects of

influenza infection itself and possible drug-disease interactions. A well-documented relationship exists between NPAEs and influenza infection itself. Differing age-related patterns of influenza-associated encephalopathy/encephalitis and NPAEs have been reported in Japanese children and adolescents, and also age-related differences exist in NAI prescribing patterns in Japan. Consequently, careful analysis is required to assess purported associations. It is important to point out that causal relationships between oseltamivir use and such events remain to be proven. Some analyses have indicated comparable or lower NPAEs rates in oseltamivir-treated compared to non-treated influenza patients (reviewed in (5)) and no higher rates of NPAEs have been found in hospitalised infants in the USA (6). Oseltamivir administration to those with influenza-associated NPAEs does not appear to worsen manifestations (7;8). Of note, the crude reporting rates for possible oseltamivir-associated NPAEs in Japan and USA were significantly lower during the 2009 pandemic than during preceding influenza seasons (9).

As pointed out by Jefferson and colleagues, the possibility of late-onset adverse events requires that sufficient follow-up be incorporated into study design to examine both possible adverse and beneficial effects. However, the low frequencies of such events would likely require much larger numbers of subjects than enrolled in most RCTs. One approach is retrospective examination of large databases that link healthcare visits, clinical diagnoses, and drug administration registries. For example, one cohort study involving over 150,000 subjects (49,238 oseltamivir recipients, 102,692 control patients) reported that oseltamivir treatment of presumed influenza was associated with lower risk of TIA or stroke in the subsequent six months (10). This kind of observational study approach has been undertaken for investigation of outcomes and possible adverse events following influenza immunisation and should also be extended to antivirals.

4. Data from observational studies. Jefferson and colleagues indicate that possible survivor treatment selection bias in observational studies can occur because patients who die early are not given the opportunity to receive treatment. However, there is also the opposite concern that sicker patients, especially in a rapidly evolving illness like influenza, are more likely to initiate therapy at any given time after symptom onset than less ill ones. This would be a conservative bias and reduce the likelihood of observing a treatment effect. Clinical experience during the 2009 H1N1 pandemic indicated that late NAI treatment in critically ill or non-surviving influenza patients was frequently due to delayed consideration of the diagnosis or failure to appreciate the potential value of starting treatment beyond two days after symptom onset in those with progressive illness or high-risk conditions. This occurred often despite some of these patients having had prior outpatient contact for their acute illness. Although the published reports indicate that most critically ill patients ultimately received antiviral therapy, delayed treatment commonly led to initiation of NAI administration as part of a salvage effort in a deteriorating patient. In part because of critical care support, even those patients who died in hospital usually survived into the second week of illness or later. Those analyzing the large amount of observational data that has been generated in recent years, particularly in the context of the 2009 H1N1 pandemic, need to keep these clinical observations in mind. Of note, a recent analysis of critically ill pandemic H1N1 patients in California compared mortality in untreated patients who survived at least to the day after symptom onset when NAIs were first given to the NAI-treated ones and found that cases who received NAI up to 4 days after symptom onset were more likely to survive ($P < 0.05$ for each day 0-4) (11).

An independent report on the observational studies of influenza antivirals published up to November 2010 (12) conducted a meta-analyses of the few studies providing effects adjusted for confounders and, while acknowledging the low quality of the evidence based on the GRADE assessment approach, concluded that in high-risk populations, oral oseltamivir may reduce mortality (odds ratio, 0.23 [95% CI, 0.13 to 0.43]) and hospitalisation (odds ratio, 0.75 [95% CI, 0.66 to 0.89]). In addition, as reported in multiple studies of hospitalised pandemic 2009 A(H1N1) patients, including high-risk ones like pregnant women and those admitted with pneumonia, treatment with oseltamivir up to 4 days and in some studies later after illness onset has been associated consistently with better outcomes (11;13-21). Such observations have served to reinforce US CDC recommendations for using influenza antivirals as early as possible in those with severe or progressive illness, those hospitalised with suspected or proven influenza, and outpatients at higher risk for influenza complications (22). Furthermore, given that the circulating influenza viruses have continued to change, with the pre-2009 A(H1N1) seasonal viruses being entirely replaced by A(H1N1)pdm09 and now antigenically drifted A(H3N2) and B viruses, ignoring observational data means that only information concerning NAI treatment for influenza viruses that are now no longer circulating is being considered.

5. Other treatment endpoints of interest. The possibility that oseltamivir might have non-specific antipyretic or immunomodulatory actions unrelated to its antiviral effects has been raised in part on the basis of murine studies (23;24). These possibilities or other symptom- modifying effects could be addressed by comparison of the course of fever and individual symptom resolution between oseltamivir and placebo recipients for those enrolled in the RTCs who did not have laboratory evidence for influenza (ITT-non-infected). Of note, antipyretics were provided to participants in these trials, so that use of paracetamol (acetaminophen) needs to be included as a confounder in such analyses.

In the published pivotal RCTs of oseltamivir treatment in adults, the fever and symptom reductions observed in oseltamivir recipients were in addition to the effects of paracetamol (acetaminophen). One previous RCT in adults with uncomplicated influenza compared amantadine to aspirin and found faster fever resolution in aspirin recipients but slower resolution of other symptoms and higher rates of adverse effects leading to drug cessation (25). While fever resolution is an objective endpoint of interest, it is generally short-lived and of limited clinical importance relative to other endpoints like time to symptom alleviation, time to return to usual activities/premorbid status, and complications reductions.

The comment by Jefferson and colleagues on measuring viral loads is confusing. Virologic endpoints like quantitative virus titers (infectious and in recent studies viral RNA), time to culture negativity, and changes in titers over time are essential to determining whether a putative influenza antiviral treatment is exerting an antiviral effect and the magnitude of that effect. Failure to detect an antiviral effect raises questions about issues like compliance, drug absorption and disposition, lack of potency, and resistance emergence. Examining such virologic measures also serves to confirm the likely mechanism of antiviral action of NAIs, inhibiting release from infected cells and spread in respiratory tract secretions to initiate subsequent rounds of replication. Several observational studies during the 2009 pandemic found that early antiviral treatment (<2-3 days from symptom onset) was associated with reduced duration of viral RNA detection (26-28). Consequently, in the context of the oseltamivir RCTs, it would be valuable to examine the correspondence between upper respiratory tract influenza virologic measures and symptom resolution and LRT complications in both oseltamivir and placebo groups.

7. Prophylaxis endpoints of interest. As indicated in my initial letter, the key efficacy endpoint for an influenza antiviral used for prophylaxis should be symptomatic, laboratory-confirmed influenza illness. Given the potential for other respiratory viruses to cause febrile respiratory illness, a focus on ILI as the primary endpoint will inevitably underestimate the protective effects of an influenza-specific chemoprophylactic agent. Of note, various definitions of symptomatic illness and ILI have been used in the influenza prophylaxis RCTs to date, so that further analyses using standardized definitions would be a helpful contribution. Other secondary endpoints of interest include laboratory documented infection (irrespective of symptoms), ILI, virus-positive ILI, and laboratory-confirmed illnesses not meeting the ILI definition. Laboratory confirmation based on both viral culture and in future studies viral RNA detection would take advantage of the greater sensitivity of RNA detection.

8. Adverse effects with prophylaxis. As detailed in the oseltamivir seasonal prophylaxis study protocols and report, the relationship between drug receipt and adverse events, including headache, in these trials (29) was determined by the study staff and investigators during the trial under blinded conditions before data lock. The assessment of causality in adverse events (unrelated, remote, possible, probable) as related to drug administration was made using pre-specified criteria in the protocol (see Appendix 2) on an individual basis by both interviewing the affected participant and considering various factors including past patterns of headaches, associated symptoms, duration and severity, timing in relation to study drug, and whether the symptom persisted during drug administration. Because of its background frequency in the population, headache is a very common event in longer term studies. When it is mild or transient despite continued drug administration, or when it occurs in context of other events (URI, trauma, stress), headache is unlikely to be drug-related. Using these criteria and the analysis report provided by the sponsor Roche, we observed headache (not otherwise specified, NOS) that was probably, possibly, or remotely related to study drug administration in 22.4% of placebo, 23.8% of once daily oseltamivir, and 25.4% of twice daily oseltamivir recipients during the 6 weeks of prophylaxis (29). The proportions were 10.2%, 8.7%, and 10.8%, respectively, for headache (NOS) that was possibly or probably related to study drug administration.

Headache is a good example of where it is essential to examine not only the frequencies of reported adverse events but also their severity and functional impact, including premature cessation of study drug. In our 6-week prophylaxis trial (29), severe headache (NOS) irrespective of relationship to study drug administration was reported in 5.0% of placebo, 3.3% of once daily oseltamivir, and 6.9% of twice daily oseltamivir, respectively. Overall premature study withdrawals were found in 21 (4.4%) of placebo, 17 (3.3%) of once daily oseltamivir, and 16 (3.1%) of twice daily oseltamivir recipients. In three placebo but no oseltamivir recipients, headache was listed as a contributory factor. However, headache was reported to be a factor leading to cessation of oseltamivir prophylaxis in one subject in another prophylaxis study (30) and was also reported at a higher frequency during 6-weeks prophylaxis in a nursing home-based RCT (5.5% placebo vs 8.3% oseltamivir)(31), so that further analyses are warranted.

9. Peer review. I thank Jefferson and his colleagues for their clarifications on the Cochrane peer review process, and as indicated above, I have provided my own suggestions on the design of future analyses by them and others. In addition, I have provided a list to the Cochrane Editorial Unit of several dozen potential expert reviewers for future protocols and reports on influenza antivirals.

Thank you for the opportunity to provide these responses and comments.

Sincerely,

Reference List

(1) Ohmit SE, Petrie JG, Cross RT, Johnson E, Monto AS. Influenza hemagglutination-inhibition antibody titer as a correlate of vaccine-induced protection. J Infect Dis 2011 Dec 15;204(12):1879-85.
(2) Hernan MA, Lipsitch M. Oseltamivir and risk of lower respiratory tract complications in patients with flu symptoms: a meta-analysis of eleven randomised clinical trials. Clin Infect Dis 2011 Aug 1;53(3):277-9.
(3) Kaiser L, Wat C, Mills T, Mahoney P, Ward P, Hayden F. Impact of oseltamivir treatment on influenza-related lower respiratory tract complications and hospitalizations. Arch Intern Med 2003 Jul 28;163(14):1667-72.
(4) Kaiser L, Keene ON, Hammond JM, Elliott M, Hayden FG. Impact of zanamivir on antibiotic use for respiratory events following acute influenza in adolescents and adults. Arch Intern Med 2000 Nov 27;160(21):3234-40.
(5) Okamoto E. Is oseltamivir (Tamiflu) safe? Re-examining the Tamiflu 'ado' from Japan. Expert Rev Pharmacoecon Outcomes Res 2010 Feb;10(1):17-24.
(6) Kimberlin DW, Shalabi M, Abzug MJ, Lang D, Jacobs RF, Storch G, et al. Safety of oseltamivir compared with the adamantanes in children less than 12 months of age. Pediatr Infect Dis J 2010 Mar;29(3):195-8.
(7) Huang YC, Li WC, Tsao KC, Huang CG, Chiu CH, Lin TY. Influenza-associated central nervous system dysfunction in Taiwanese children: clinical characteristics and outcomes with and without administration of oseltamivir. Pediatr Infect Dis J 2009 Jul;28(7):647-8.
(8) Tanabe T, Hara K, Nakajima M, Shimakawa S, Tamai H. Oseltamivir treatment for children showing abnormal behavior during influenza virus infection. Brain Dev 2010 Jun;32(6):440-4.
(9) Donner B, Bader-Weder S, Schwarz R, Peng MM, Smith JR, Niranjan V. Safety profile of oseltamivir during the 2009 influenza pandemic. Pharmacoepidemiol Drug Saf 2011 May;20(5):532-43.
(10) Madjid M, Curkendall S, Blumentals WA. The influence of oseltamivir treatment on the risk of stroke after influenza infection. Cardiology 2009;113(2):98-107.
(11) Louie JK, Yang S, Acosta M, Yen C, Samuel MC, Schechter R, et al. Treatment with neuraminidase inhibitors for critically ill patients with influenza A(H1N1)pdm09. Clin Infect Dis 2012 Jul 26.
(12) Hsu J, Santesso N, Mustafa R, Brozek J, Chen YL, Hopkins JP, et al. Antivirals for treatment of influenza: a systematic review and meta-analysis of observational studies. Ann Intern Med 2012 Apr 3;156(7):512-24.
(13) Yang SG, Cao B, Liang LR, Li XL, Xiao YH, Cao ZX, et al. Antiviral therapy and outcomes of patients with

pneumonia caused by influenza A pandemic (H1N1) virus. PLoS One 2012;7(1):e29652.
(14) Higuera Iglesias AL, Kudo K, Manabe T, Corcho Berdugo AE, Corrales BA, Alfaro RL, et al. Reducing occurrence and severity of pneumonia due to pandemic H1N1 2009 by early oseltamivir administration: a retrospective study in Mexico. PLoS One 2011;6(7):e21838.
(15) Mosby LG, Rasmussen SA, Jamieson DJ. 2009 pandemic influenza A (H1N1) in pregnancy: a systematic review of the literature. Am J Obstet Gynecol 2011 Jul;205(1):10-8.
(16) Dubar G, Azria E, Tesniere A, Dupont H, Le RC, Baugnon T, et al. French experience of 2009 A/H1N1v influenza in pregnant women. PLoS One 2010;5(10).
(17) Rodriguez A, Zaragoza R, Daz E, Daz JJ, Marques A. Early oseltamivir treatment was associated with improved outcomes in 2009 pandemic influenza A (H1N1)v in Spain. Intensive Care Med 2010;36:S136.
(18) Siston AM, Rasmussen SA, Honein MA, Fry AM, Seib K, Callaghan WM, et al. Pandemic 2009 influenza A(H1N1) virus illness among pregnant women in the United States. JAMA 2010 Apr 21;303(15):1517-25.
(19) Sugaya N, Shinjoh M, Mitamura K, Takahashi T. Very low pandemic influenza A (H1N1) 2009 mortality associated with early neuraminidase inhibitor treatment in Japan: analysis of 1000 hospitalised children. J Infect 2011 Oct;63(4):288-94.
(20) Yu H, Feng Z, Uyeki TM, Liao Q, Zhou L, Feng L, et al. Risk factors for severe illness with 2009 pandemic influenza A (H1N1) virus infection in China. Clin Infect Dis 2011 Feb 15;52(4):457-65.
(21) Viasus D, Pano-Pardo JR, Pachon J, Riera M, Lopez-Medrano F, Payeras A, et al. Timing of oseltamivir administration and outcomes in hospitalised adults with pandemic 2009 influenza A(H1N1) virus infection. Chest 2011 Oct;140(4):1025-32.
(22) "Have you heard?": CDC recommendations for influenza antiviral medications remain unchanged. http://www.cdc.gov/media/haveyouheard/stories/Influenza_antiviral.html. 2012.
(23) Ono H, Nagano Y, Matsunami N, Sugiyama S, Yamamoto S, Tanabe M. Oseltamivir, an anti-influenza virus drug, produces hypothermia in mice. Biol Pharm Bull 2008 Apr;31(4):638-42.
(24) Wong ZX, Jones JE, Anderson GP, Gualano RC. Oseltamivir treatment of mice before or after mild influenza infection reduced cellular and cytokine inflammation in the lung. Influenza Other Respi Viruses 2011 Sep;5(5):343-50.
(25) Younkin SW, Betts RF, Roth FK, Douglas RG, Jr. Reduction in fever and symptoms in young adults with influenza A/Brazil/78 H1N1 infection after treatment with aspirin or amantadine. Antimicrob Agents Chemother 1983 Apr;23(4):577-82.
(26) Cao B, Li XW, Mao Y, Wang J, Lu HZ, Chen YS, et al. Clinical features of the initial cases of 2009 pandemic influenza A (H1N1) virus infection in China. N Engl J Med 2009 Dec 24;361(26):2507-17.
(27) Ling LM, Chow AL, Lye DC, Tan AS, Krishnan P, Cui L, et al. Effects of early oseltamivir therapy on viral shedding in 2009 pandemic influenza A (H1N1) virus infection. Clin Infect Dis 2010 Apr 1;50(7):963-9.
(28) Li IW, Hung IF, To KK, Chan KH, Wong SS, Chan JF, et al. The natural viral load profile of patients with pandemic 2009 influenza A(H1N1) and the effect of oseltamivir treatment. Chest 2010 Apr;137(4):759-68.
(29) Hayden FG, Atmar RL, Schilling M, Johnson C, Poretz D, Paar D, et al. Use of the selective oral neuraminidase inhibitor oseltamivir to prevent influenza. N Engl J Med 1999 Oct 28;341(18):1336-43.
(30) Welliver R, Monto AS, Carewicz O, Schatteman E, Hassman M, Hedrick J, et al. Effectiveness of oseltamivir in preventing influenza in household contacts: a randomised controlled trial. JAMA 2001 Feb 14;285(6):748-54.
(31) Peters PH, Jr., Gravenstein S, Norwood P, De B, V, Van CA, Gibbens M, et al. Long-term use of oseltamivir for the prophylaxis of influenza in a vaccinated frail older population. J Am Geriatr Soc 2001 Aug;49(8):1025-31.

Appendix 2 Definition of Adverse Event Relationship to Treatment

Probable

This category applies to those adverse events which are considered, with a high degree of certainty, to be related to the test drug. An adverse event may be considered probable if:

1. It follows a reasonable temporal sequence from administration of the study drug.
2. It cannot be reasonably explained by the known characteristics of the subject's clinical state, environmental or toxic factors, or other modes of therapy administered to the subject.
3. It disappears or decreases on cessation or reduction of dose. (There are important exceptions when an adverse event does not disappear upon discontinuation of the drug, yet drug- relatedness clearly exists; e.g., (1) bone marrow depression, (2) tardive dyskinesias).
4. It follows a known pattern of response to the study drug.
5. It reappears upon re-challenge.

Possible

This category applies to those adverse events in which the connection with the test drug administration appears unlikely but cannot be ruled out with certainty. An adverse event may be considered possible if or when:

1. It follows a reasonable temporal sequence from the administration of study drug.
2. It may have been produced by the subject's clinical state, environmental or toxic factors, or other modes of therapy administered to the subject.
3. It follows a known pattern of response to the study drug.

Remote

In general, this category is applicable to an adverse event which meets the following criteria:

1. It does not follow a reasonable temporal sequence from administration of the study drug.
2. It may readily have been produced by the subject's clinical state, environmental or toxic factors, or other modes of therapy administered to the subject.
3. It does not follow a known pattern of response to the study drug.
4. It does not reappear or worsen when the drug is re-administered.

Unrelated

This category is applicable to those adverse events which are judged to be clearly and incontrovertibly due only to extraneous causes (disease, environment, etc.) and do not meet the criteria for drug relationship listed under remote,possible, or probable.

	Probable	Possible	Remote	Unrelated
Clearly due to extraneous causes	-	-	-	+
Reasonable temporal association with drug administration	+	+	-	-
May be produced by subjects clinical state	-	+	+	+
Known response pattern to suspected drug	+	+	-	-
Disappears or decreases on cessation or reduction in dose	+	-	-	-
Reappears on re-challenge	+	-	-	-

Reply

Reply to [redacted] 10 August 2012

Thank you for taking the trouble to provide further feedback to our responses to your first set of feedback comments.

You remain concerned about 1) "…selective approach to data analysis and presentation…", especially with respect to our concern that ITT-infected (ITTI) criteria are inappropriate; and 2) our identification of biases that may exaggerate the effectiveness of oseltamivir. You detail these concerns in more detail:

1. ITT and ITTI

You propose an analysis of ITTI in which patients are categorised not by an immune response (which we regard

as potentially flawed because our interpretation of the data suggests the drug may interfere with the immune response), but instead by determining whether patients were seroconverting excreting influenza virus at enrolment.

This sounds sensible, and were the data of symptoms and baseline infectivity (by serology or even virus shedding) available to us in suitable format, we would include this analysis. By this, we would expect the randomisation of patients into the two groups to be independent of the initiation of the drug (that is the "influenza-positive" or "-negative") before the drug was administered, in case (as may be with the immune response) the drug interferes with virus excretion (as the manufacturer claims in some of its literature).

You also propose an analysis of those grouped by ITTI from serological conversion with those grouped by virus excretion. This also would be useful, to determine whether or not a bias exists in the current data (in either direction, as you point out – the possible mechanisms you outline are plausible).

However, your hypothesis "If oseltamivir is most beneficial in preventing lower respiratory tract (LRT) complications" IS one of the main issues to be confirmed.

As already described in our review, you reported a reduction of cytokine production in response to influenza infection by oseltamivir in humans:

- Hayden FG, Atmar RL, Schilling M, Johnson C, Poretz D, Paar D, et al. Use of the selective oral neuraminidase inhibitor oseltamivir to prevent influenza. New England Journal of Medicine 1999;341(18):1336-43

These findings suggest that reduction of antibody production cannot simply be assumed to be the result of reduced viral load.

2. Sample sizes
You describe in more detail the Kaiser 2003 pooled analysis of complications:

- Kaiser L, Wat C, Mills T, Mahoney P, Ward P, Hayden F. Impact of oseltamivir treatment on influenza-related lower respiratory tract complications and hospitalizations. Arch Intern Med 2003;163:1667-72

This was central to the start of our unease, after it was pointed out to us (in this Feedback section!) by Hayashi that over half of the data in it were of unpublished trials. You state that the end-points were established a priori and not post hoc. You admit to shortcomings of the paper, but point out that they were declared in the paper itself. You suggest that because the two published trials meta-analysed had no more favourable drug results than the unpublished, bias is less likely.

We think this is to misunderstand our central concern: we are unable to critically appraise the trials in the usual way because they are not available to us, nor, apparently, any other group unselected by the manufacturer. Incidentally we note that you yourself, even as an author, admit you were unable to locate the data for this paper on request, referring us instead to the sponsoring manufacturer, Roche:

- Cohen D. Complications: tracking down the data on oseltamivir. BMJ 2009;339:b5387.

This inability by you (authors) or sponsoring manufacturer to provide data for independent scrutiny is disgraceful, a view shared by others, http://bmj.com/tamiflu.

3. Adverse effects of NIs
We find it interesting that you call these adverse events 'complications'. You point to our concerns about neuropsychiatric adverse events (NPAEs), and (correctly) state that any association recorded in the literature "…remains to be proven…" with some references (all were retrospective studies and mostly sponsored by the manufacturer) that suggest that there is no increase over control groups. We have other references suggesting the opposite:

- Hama.R. Fatal neuropsychiatric adverse reactions to oseltamivir: case series and overview of causal relationship. Int J Risk Safety Med: 20 (2008): 5-36 : http://npojip.org/english/no11.html

- Nakamura K, Schwartz BS, Lindegårdh N, Keh C, Guglielmo BJ. Possible neuropsychiatric reaction to high-dose oseltamivir during acute 2009 H1N1 influenza A infection. Clin Infect Dis. 2010 Apr 1;50:e47-9.

- Kruker AT, Krause M. ["Oseltamivir-induced delirium"]. Ther Umsch. 2010 Dec;67(12):613-5. German.

- Chung S, Joung YS. Oseltamivir (tamiflu) induced depressive episode in a female adolescent. Psychiatry Investig. 2010 Dec;7(4):302-4. Epub 2010 Nov 11.

The following are prospective cohort studies that aimed to analyze the association of NPAEs and administration of NIs, in particular oseltamivir.

- F. Fujiwara, S. Ikushima, N. Hibi et al. An analysis of Risk factors of abnormal behavior in two seasons (07, 08) of influenza infection. presentation at the 40th annual meeting of the Japanese Society for paediatric Infectious Diseases held on 15 and 16 (2008)

- Fujita T. Fujii Y, Watanabe Y, Mori M, Yokota S. A Pharmacoepidemiological Study on the Relationship between Neuropsychiatric Symptoms and Therapeutic Drugs after Influenza Infection. Jap J Pharmacoepidemiol 2010; 15: 73-92.

This preliminary report on the analysis of randomised controlled trials of oseltamivir for prophylaxis contains our response to Roche's report discussing NPAEs and oseltamivir:

- Jones, M., Hama, R., Jefferson, T., Doshi, P. Neuropsychiatric adverse events and Oseltamivir for prophylaxis (letter). Drug Safety, 2012, 35 (12): 1187-1190.

A proportional mortality study indicates that oseltamivir increases sudden death (odds ratio: 5.9) compared with zanamivir users in an analysis of all deaths among ˜ 20 million 2009A/H1N1 influenza patients in Japan. This effect is also observed for the comparison of oseltamivir users with non-users.

- Hama R, Jones M, Okushima H, Kitao M, Noda N, Hayashi K, Sakaguchi K. Oseltamivir and early deterioration leading to death: a proportional mortality study for 2009A/H1N1 influenza. Int J Risk Saf Med. 2011;23(4):201-15. http://iospress.metapress.com/content/5257410g24403m68/fulltext.pdf

We have presented many of these studies in our previous reply to you, without response.

Of course the uncertainty about causation is true for many drug adverse events: our duty is to ensure that any such uncertainty is clearly articulated.

Nevertheless we entirely agree that "…observational studies … undertaken for investigation of outcomes and possible adverse events following influenza immunisation … should also be extended to antivirals." However, because this Cochrane review is limited to randomised data, such observational studies would be conducted outside this particular review.

4. Observational data
You point to our concerns about observational data in general for answering intervention questions. We acknowledge the plethora of observational data available, and even the meta-analysis of some of them. This does not detract from our continued concern that the best data for answering these questions are randomised, and to leave most of these data unavailable for independent scrutiny is unforgivable.

Moreover, the observational studies are regarded as poor in quality. A recent systematic review and meta-analysis of observational data for antivirals for the treatment of influenza concluded, "…therapy with oral oseltamivir and inhaled zanamivir may provide a net benefit over no treatment of influenza. However the confidence in the estimates of the effects for decision making is low to very low."

- Hsu J, Santesso N, Mustafa R, Brozek J, Chen YL, Hopkins JP, Cheung A, Hovhannisyan G, Ivanova L, Flottorp SA, Saeterdal I, Wong AD, Tian J, Uyeki TM, Akl EA, Alonso-Coello P, Smaill F, Schünemann HJ. Ann Intern Med. 2012 Apr 3;156(7):512-24. doi: 10.1059/0003-4819-156-7-201204030-00411. Epub 2012 Feb 27. Antivirals for treatment of influenza: a systematic review and meta-analysis of observational studies

Incidentally, we are interested in rigorously meta-analysing these data ourselves, and have put in a protocol to do just that. (Jones M, Hama R. Effect of oseltamivir on mortality in treatment of 2009A/H1N1 influenza patients. PROSPERO 2012:CRD42012002245 Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42012002245

The proportional mortality study (above), analysing all influenza deaths in Japan and estimating populations who took antivirals and did not take them as the denominators, provides far more reliable estimates of risk from drug exposures than retrospective analysis of surveillance cases without exposed populations (denominators). Contrary to your suggestion "…there is also the opposite concern that sicker patients, especially in a rapidly evolving illness like influenza, are more likely to initiate therapy at any given time after symptom onset than less ill ones…", no such tendency was detected in this study. Proportions of patients treated with antivirals within 12 hours from the onset of fever were significantly lower in the "not mild" cases (26.5%) than "mild" cases (35.4%) at the time when antiviral was prescribed [Table 2b]. However, no patients who deteriorated before the first presentation at medical facilities were treated with antivirals before deterioration [Table 2a], while 78% of "mild" cases and 55% of "not mild" cases were prescribed antivirals within 48 hours from onset of fever [Tables 2a and 2b]. These may be related to the lower positive results (45%) of rapid testing for influenza virus in the "not mild" cases than that in the "mild" cases (60%) at the first consultation:

- Hama R, Jones M, Okushima H, Kitao M, Noda N, Hayashi K, Sakaguchi K. Oseltamivir and early deterioration leading to death: a proportional mortality study for 2009A/H1N1 influenza. Int J Risk Saf Med. 2011;23(4):201-15. http://iospress.metapress.com/content/5257410g24403m68/fulltext.pdf

5. Other treatment endpoints of interest
Does oseltamivir have non-specific antipyretic or immune-modulatory actions unrelated to its antiviral effect?

We have already noted the hypothermic and immune-suppression effect of oseltamivir in humans, some from your own writing.

- Hama.R. Fatal neuropsychiatric adverse reactions to oseltamivir: case series and overview of causal relationship. Int J Risk Safety Med 2008:20:5-36
- Hayden FG, Treanor JJ, Fritz RS, Lobo M, Betts RF, Miller M, et al. Use of the oral neuraminidase inhibitor oseltamivir in experimental human influenza: randomised controlled trials for prevention and treatment. JAMA 1999;282:1240-6.

Your suggestion that antipyretic actions of oseltamivir be tested by comparing those randomised to oseltamivir against those not in the non-ITTI group is worth consideration, (although the results might be difficult to interpret). Again, as mentioned above, it would be good to have access to sufficient data to allow this analysis and others we have outlined in the Protocol.

We note your criticism about over-focusing on fever as a proxy for symptom resolution. We are of course interested in any good measure of the latter that is not only objective, but also common to all trials. Nevertheless, despite your criticism, fever is a reasonable marker of 'illness' from infections such as influenza, and probably correlates reasonably well with symptom resolution (especially in the prophylaxis trials) and in

the treatment trials (if fever is measured until complete resolution) – it is, after all, a cardinal symptom – and has the great advantage of being clearly measured.

You suggest that we test whether viral excretion correlates with symptoms of influenza. We agree that this would be an interesting analysis, were the data available to us, (see above).

7. (Note there was no Point 6) Should we be focusing so much on influenza-like illness (ILI)?

Of course, if oseltamivir neither reduces antibody production to influenza virus nor conceals testing positivity, selecting only laboratory-confirmed influenza might be a reasonable end point for prophylaxis trials. However the facts suggest these cannot be assumed.

In any case, the Cochrane Collaboration is dedicated to finding the best available evidence to enable patients and their clinicians to make best-informed decisions. To that end, ILI is what the vast majority of clinicians and their patients will be facing. Therefore this is an end-point of direct relevance to them, and we make no apology for including it.

8. Adverse events in prophylactic trials
Thanks for this detailed information. Further analyses are indeed what we would like to undertake according to our protocol.

9. Peer review
Thanks for offering a list of your own colleagues to act as peer reviewers. We adhere to the principle of ensuring there is methodological expertise as well as content expertise. Your list will be useful to consider when finding peer reviewers.

As you may be aware, because this particular Review Group (Acute Respiratory Infections) has its Co-ordinating Editor as an Author on this review, the handling of the manuscript is managed by the Central Editorial Unit to minimise any potential conflict of interest.

Contributors

Chris Del Mar, Tom Jefferson, Rokuro Hama, Mark Jones, Peter Doshi, Carl Heneghan, Matthew Thomson.

Feedback from [redacted], 13 February 2013

Summary

Comment: The selection criteria in the review seem highly unusual. The authors describe a 2-stage process for including trials.

In the first stage, they require that the trial reports they analyse have "external consistency". As far as I can tell, this means that they must be able to verify the contents of the report from an external source.

This seems an extraordinarily high bar to set. I am not aware that it is part of standard Cochrane methodology. If it were applied across Cochrane reviews more generally, I imagine that very few Cochrane reviews would include any evidence at all, especially given that most Cochrane reviews are done perfectly happily with

published papers, whereas this one had the advantage of clinical study reports, which are generally far more reliable and comprehensive than published papers.

It is almost as if the authors have gone out of their way to exclude the evidence, which does not help to answer important questions about the efficacy of neuraminidase inhibitors.

It is also noteworthy that no specific reasons were given for exclusion of studies from stage I of the process: we are only told that "insufficient information was available". In the interests of transparency, it would be better to know specifically what information was lacking.

May I suggest that the authors either explain the reason why they felt the need to use far stricter inclusion criteria than is normal in Cochrane reviews, or revisit their inclusion criteria so that the studies can be analysed.

I agree with the conflict of interest statement below:

I certify that I have no affiliations with or involvement in any organization or entity with a financial interest in the subject matter of my feedback.

[redacted] Dianthus Medical Limited

Reply

[redacted] writes:

“The selection criteria in the review seem highly unusual. The authors describe a 2-stage process for including trials. In the first stage, they require that the trial reports they analyse have "external consistency". As far as I can tell, this means that they must be able to verify the contents of the report from an external source.”

At page 11 of the review we provide the definition: “External consistency. Consistency of data as reported in regulatory documents, other versions of the same clinical study reports/unpublished reports and other references, to be established by cross-checking”

“This seems an extraordinarily high bar to set. I am not aware that it is part of standard Cochrane methodology. If it were applied across Cochrane reviews more generally, I imagine that very few Cochrane reviews would include any evidence at all, especially given that most Cochrane reviews are done perfectly happily with published papers, whereas this one had the advantage of clinical study reports, which are generally far more reliable and comprehensive than published papers”.

And

“May I suggest that the authors either explain the reason why they felt the need to use far stricter inclusion criteria than is normal in Cochrane reviews, or revisit their inclusion criteria so that the studies can be analysed.”

Our review is the first systematic review that we are aware of to be completely based on regulatory information. As our basic element of data synthesis was different, we had to develop new methods which we did transparently and are described in the review. It was a fact that we had received partial clinical study reports for the same trials from both Roche and EMA. We felt the need to ensure these reports were consistent. Whether our methods were an “extraordinarily high bar” or a reasonable bar or too low a bar is a judgment readers can make for themselves.

The background history which informed our methodology is explained in the review itself. At pages 4 and 5 of the review we write:

“In 2009, a reader posted a comment in response to the (then current) 2006 version of this review (Jefferson 2006). He pointed out that the review had endorsed the claim regarding a reduction in complications based on the uncritical inclusion of the Kaiser meta-analysis (Doshi 2009). The reader pointed out that only two of the 10 ’Kaiser trials’ had been published (Nicholson 2000; Treanor 2000) and the information provided by the Kaiser text about the remaining eight was insufficient for their appraisal. Our subsequent efforts to retrieve and review the eight unpublished trials (representing 2691 patients) were unsuccessful, raising the possibility that the findings of our previous review were not an accurate estimate of the benefits and safety of the drug. In addition, we found clear evidence of possible publication bias (see below) amid concern that some evaluations have not been available to scrutiny by the scientific community (Cohen 2009; Doshi 2009; Freemantle 2009; Godlee 2009).”

“This review is focused on healthy adults and children. It represents the amalgamation of two long-standing Cochrane reviews on the effects of NIs for influenza in healthy adults (Jefferson 2010a, also published as Jefferson 2009a) and children (Matheson 2007). The reviews were combined to pool our collective expertise and time in extracting and assessing data from clinical study reports, which in the case of some oseltamivir trials, report both adult and paediatric outcomes. Cochrane reviews of NIs in both children and adults generated intense interest from clinicians and media during the influenza outbreak declared a pandemic by the WHO in 2009. The Cochrane review of NIs in healthy adults highlighted the high risk of publication bias (Jefferson 2010a). In 2009, a reader posted a comment in response to the (then current) 2006 version of this review (Jefferson 2006). He pointed out that the review had endorsed the claim regarding a reduction in complications based on the uncritical inclusion of the Kaiser meta-analysis (Doshi 2009). The reader pointed out that only two of the 10 ’Kaiser trials’ had been published (Nicholson 2000; Treanor 2000) and the information provided by the Kaiser text about the remaining eight was insufficient for their appraisal. Our subsequent efforts to retrieve and review the eight unpublished trials (representing 2691 patients) were unsuccessful, raising the possibility that the findings of our previous review were not an accurate estimate of the benefits and safety of the drug. In addition, we found clear evidence of possible publication bias (see below) amid concern that some evaluations have not been available to scrutiny by the scientific community (Cohen 2009; Doshi 2009; Freemantle 2009; Godlee 2009).

Our attempts to reconcile published and unpublished evidence by contacting the manufacturer and study authors failed (the latter were unable to provide us with the necessary data; some were not in possession of the data and others may have been restricted by confidentiality agreements). Together with the British Medical Journal (BMJ)we ascertained that ghostwriters had been involved, which means the named authors may not have been in full control of the trial publications (Cohen 2009). We also identified several key differences in licensed indications for oseltamivir between regulatory systems (mainly between the US, Europe and Japan) and under-reporting of harms. The differences are detailed elsewhere (Doshi 2009) but of particular concern was the insistence of the FDA that oseltamivir has not been shown to reduce complications (FDA 2011a). The FDA has also not allowed an indication for interference of viral transmission within households (the key concept behind post-exposure prophylaxis). This undermined our confidence in published data and in the findings of our previous Cochrane reviews. In the background of all this were suggestions that NIs may not be as safe as previously assumed, with associations between oseltamivir use and neuropsychiatric adverse reactions of particular concern (Hama 2008).”

[redacted] writes:

“It is almost as if the authors have gone out of their way to exclude the evidence, which does not help to answer important questions about the efficacy of neuraminidase inhibitors."

A page 5 of the review we write:

“During the preparation of the 2010 review and of the current review, we realised that there were multiple sources and different levels of granularity of clinical trial data (see ’The Scope of Clinical Trial Data’ table in

Jefferson 2011).We decided that clinical study reports and regulatory comments were likely to provide the least biased, most complete and most insightful set of data for our review".

And

"We identified that 60% (3145/5267) of patient data from randomised, placebo-controlled phase III treatment trials of oseltamivir have never been published. This includes M76001, the biggest treatment trial ever undertaken on oseltamivir (with just over 1400 people of all ages). Exclusion of unpublished data changed our previous findings regarding oseltamivir's ability to reduce complications of influenza (Doshi 2009; Jefferson 2009a)."

Our attempts at identifying and retrieving all available evidence from regulators and manufacturers since 2009 are documented at http://bmj.com/tamiflu.

[redacted] writes:

"It is also noteworthy that no specific reasons were given for exclusion of studies from stage I of the process: we are only told that "insufficient information was available". In the interests of transparency, it would be better to know specifically what information was lacking."

In Table 9 (page 186) we list all studies included in Stage 1 and report details of what data for each were available to us. For, example for trial MV22940 we know that it is likely to be a randomised trial assessing effects of oseltamivir on post exposure prophylaxis but no other data are available to us. In these circumstances we cannot proceed to assessment until the information is available, as explained in the text of the review. However these studies are not excluded but are marked as pending assessment.

We invite Adam Jacobs to read the review and the references which document the history of the review, background and rationale for withdrawing the original review and developing the current version. We also invite [redacted] to clarify what business relation his firm has if any with Roche, GSK and BioCryst Ltd.

It is possible that future Cochrane reviews will include an increasing proportion of regulatory information to minimize the effects of reporting bias. This type of speculation is however beyond the scope of the review.

Contributors

Cochrane Neuraminidase Inhibitors Review Team, March 5, 2013

Prof Chris Del Mar, Coordinating Editor, Acute Respiratory Infections Cochrane Review Group, Australia
Dr Peter Doshi, Postdoctoral Fellow, Johns Hopkins University, USA
Dr Rokuro Hama, Physician, Pharmaco-epidemiologist, Japan Institute of Pharmaco-vigilance, University of Osaka, Japan
Dr Carl Heneghan, Clinical Reader, Department of Primary Care Health Sciences, University of Oxford, UK
Dr Tom Jefferson, Epidemiologist, Acute Respiratory Infections Cochrane Review Group, Italy
Dr Mark Jones, Statistician, University of Queensland, Australia
Dr Matthew Thompson, Clinical Reader, Department of Primary Care Health Sciences, University of Oxford, UK

Feedback from [redacted], 6 May 2013

Summary

Comment: Oseltamivir (Tamiflu) shortens the duration of influenza-like illness by 13% (95% CI: 8% to 18%)

In studies measuring dichotomous outcomes, relative risk (RR) is a standard measure for comparing study groups. The purpose of using RR is to adjust for baseline variability in the occurrence of disease. It is easier to compare two trials on the basis of their RR estimates than on the basis of their absolute effects.

The relative effect should also be calculated for continuous outcomes. Although the duration of disease may vary randomly in placebo groups, there are also biological reasons why diseases in different placebo groups differ in their severity and duration. For example, in Analysis 1.1 of this review, the duration of influenza-like illness in the placebo group of trial WV15671 is 35% shorter than in the placebo group of trial WV15819/WV15876/WV15978 (Z = 6.5; P = <0.00001; 125h/192h). Such very large baseline differences are not explained by chance. Differences in the study populations, influenza seasons, study protocols, etc. are plausible explanations for the baseline variation. The above-mentioned baseline difference is much greater than any of those between the oseltamivir (Tamiflu) and placebo groups in the five trials of Analysis 1.1. As for dichotomous outcomes, the baseline variability of continuous outcomes can be adjusted for by calculating the effect in percentages, i.e., the relative effect. Furthermore, the percentage effect is informative for an average reader because the reader may form an opinion on whether, for example, a 10% or 20% average decrease in the duration is worth the cost and effort of the treatment. Separate from the absolute effect in days, the percentage effect shows whether the effect is small or large.

Therefore the effect of oseltamivir should be calculated also as a percentage effect. I calculated the relative effects for the five trials listed in Analysis 1.1, pooled them using the fixed effect inverse variance method of RevMan, and found that the average effect of oseltamivir is a 13% (95% CI: 8 to 18%) decrease in the duration of influenza-like illness.

Furthermore, the relative effect estimate makes it possible to compare the effects of treatments for related conditions. Influenza-like illness has substantial overlap with the common cold. In our Cochrane review on vitamin C and the common cold we calculated that ≥1 g/day of vitamin C shortens colds in adults by 8% (95% CI: 4 to 12%) and in children by 18% (95% CI: 9 to 27%) [1]. Another meta-analysis found that a high dose of zinc (>75 mg/day) as zinc acetate lozenges decreased the duration of colds by 42% (95% CI: 35 to 48%) and as zinc lozenges made with other salts by 20% (95% CI: 12 to 28%)[2]. The mechanism of the effect of vitamin C and zinc lozenges is not understood; however, there is no reason to assume that their effects are specific, for example, to the rhinovirus. If vitamin C and zinc lozenges have effects on diverse respiratory viruses, they might also have an effect on influenza viruses. In mice, influenza infection decreased vitamin C concentration in bronchoalveolar lavage fluid [3]. In mice, vitamin C deficiency increased lung pathology caused by influenza infection [4]. An early study with influenza patients reported that the occurrence of pneumonia was 80% lower (2 vs. 10 cases) in the vitamin C group, suggesting that vitamin C might also have an effect on influenza in humans [5,6]. If the effects of vitamin C and zinc lozenges on influenza-like illness are of the same magnitude as their effects on the common cold, then the effects of these treatments compare reasonably with oseltamivir. The comparison of the percentage effects of oseltamivir, vitamin C and zinc lozenges may be useful when considering how future research resources concerning the treatment of respiratory virus infections might be allocated. In this respect, the type of effect measure has a much wider importance than just its use in evaluating the effectiveness of oseltamivir as an issue of its own.

Thus the relative effect estimate adjusts for baseline variations between trials, it is informative for most readers because people are familiar with percentages, and it makes it easier to compare different treatments for related conditions. For these reasons I would like to encourage the authors to calculate and report the relative effect estimates for oseltamivir in the next revision of the review.

References

[1] Hemilä H, Chalker E. Vitamin C for preventing and treating the common cold. Cochrane Database Syst Rev 2013:CD000980. http://dx.doi.org/10.1002/14651858.CD000980.pub4

[2] Hemilä H. Zinc lozenges may shorten the duration of colds: a systematic review. Open Respir Med J 2011;5:51-8. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3136969/

[3] Buffinton GD, Christen S, Peterhans E, Stocker R. Oxidative stress in lungs of mice infected with influenza A virus. Free Rad Res Commun 1992;16:99-110 http://www.ncbi.nlm.nih.gov/pubmed/1321077, http://dx.doi.org/10.3109/10715769209049163

[4] Li W, Maeda N, Beck MA. Vitamin C deficiency increases the lung pathology of influenza virus-infected gulo-/- mice. J Nutr 2006;136:2611-6. http://www.ncbi.nlm.nih.gov/pubmed/16988135, http://jn.nutrition.org/content/136/10/2611

[5] Kimbarowski JA, Mokrow NJ. Colored precipitation reaction of the urine according to Kimbarowski as an index of the effect of ascorbic acid during treatment of viral influenza [in German]. Deutsche Gesundheitswesen 1967;22:2413-8. http://www.ncbi.nlm.nih.gov/pubmed/5614915, Translation: http://www.mv.helsinki.fi/home/hemila/T4.pdf

[6] Hemilä H, Louhiala P. Vitamin C for preventing and treating pneumonia. Cochrane Database Syst Rev 2007:CD005532. http://dx.doi.org/10.1002/14651858.CD005532.pub2

I agree with the conflict of interest statement below:
I certify that I have no affiliations with or involvement in any organization or entity with a financial interest in the subject matter of my feedback.

[redacted]
Department of Public Health, University of Helsinki

Reply

Thank you for your suggestion and comprehensive argument why you think it is important. Indeed in our 2006 and 2009 updates of A047 (the previous review on antivirals for influenza in otherwise healthy adults) we pooled hazard ratios and reported relative effects for time to alleviation of symptoms. However GSK, the manufacturer of zanamivir, made the comment that hazard ratios may not be appropriate due to non-proportional hazards. Therefore for A159 we reported absolute treatment effects for time to alleviation of symptoms but not relative effects. We agree with your argument and will report absolute and relative effects for time to alleviation of symptoms and other outcomes in the next update of 'Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children' due at the end of 2013.

Contributors

Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ

Review amendments, 16 May 2013

Summary

As reported in the current version of our review, we will complete the review of regulatory information which arrived after our original timelock. We will assess additional evidence from oseltamivir Modules 2, evidence on adverse events following exposure to neuraminidase inhibitors (NIs) and clinically relevant outcomes.

A rationale and description of our methods follows.

Evidence from Modules 2 (Ms2) of oseltamivir trials

1. Summary and background

This part of the document will describe our efforts to determine whether the additional information included within Modules 2 (Ms2) of clinical study reports (CSRs) would change the risk of bias assessment, identify additional useful or relevant information, and conclusions of the overall body of evidence contained within our existing review. A second aim is to construct and test a tool that could be used to extract, organise and appraise study information contained in such modules.

The items which are most commonly found in the M2 of the oseltamivir trials are Certificates of Analysis (a report on the colour, composition and content of active and control substance capsules, blank Case Report Forms (case notes for each participant), Follow up cards/Diary cards (on which each participant recorded information such as symptoms), Informed Consent text and participant contract (to be administered to and signed by each participant), Lists of Investigators in the trial, Investigation review Board, Ethics committees and Study Sites' Addresses, the Reporting Analysis Plan (Roche's term for the Statistical Analysis Plan or SAP detailing the types of data analyses to be carried out), Randomisation List (used to allocate participants and the study Protocol with its amendments when appropriate or available.

1.2 Methods

We received 12 CSR Ms2 from 31 studies requested from EMA by July 2011. Before we reviewed Ms2 we knew they contained protocols, with their amendments, certificate of analyses, blank case report forms, randomisation and participating centres' lists. However, we had no precise idea whether this was a comprehensive list or whether further items would be identified once we started reviewing. We also noted that the same info was reported elsewhere in the CSRs (for example in the core report) but in a different level of detail. A good example of this is the statistical analysis section of the core report which is a few pages long chapter, compared to the Statistical Analysis Plan (SAP), which is a self-contained document included in M2. In addition we were not aware of the existence of any readily available tool to allow us to extract, organise and appraise the information contained in the Ms2.

As consequence we decided to develop our own tool. Our plan is to do this by identifying the types of items contained in the Ms2 available to us and their location in the Ms2. The outline content of all items identified will be checked in the Ms2 because of the potential for differing titles for the same item. For example we have already noticed that Research Analysis Plan (RAP) is sometimes called Data Analysis Plan (DAP) or Statistical Analysis Plan (SAP). Another example are the Protocol Amendment Histories and Protocol Modification History Document. These represented different ways of identifying the same item and need to be given a single identifier. Items such as Data Reporting and Analysis Manual (DRAM) are only cited in one M2. We will also conduct a pilot to identify with certainty which items are present more frequently. We will make a list of what we thought were most present and important items contained in the Ms2 and create a grid based on the sequence of development of the trial design and analysis plan. For example we want to track whether the reporting of the trial study design in the relevant section of the protocol and its amendments (in M2) is consistent with that described in the core report (in M1). We will also make an initial extraction frame to reconstruct the timeline of the study documents, summarising the number of protocol changes and their dates in sequence. This has the purpose of giving an overview of the main timeline points of the key items of study design and analysis.

We will then pilot our extraction sheet and make changes following discussion with all authors. We will extract the data in the same groups we worked in the original review.

We will define the impact of adding M2 information by measuring the change in risk of bias (ROB) assessment in our review as well as reporting our summary description and appraisal of each trial before and after addition of the data and comparing it with the manufacturer's assessment.

The detailed questions addressed by our analysis are:

1. Does addition of M2 to M1 change the risk of bias evaluation compared to M1 alone?
2. Does reading M2 and M1 in CSRs change the risk of bias evaluation compared to using published papers?
3. Is the current risk of bias tool adequate for assessing trials based on reading M2 then M1 in the CSRs?
4. Does reading M2 and M1 in the CSRs identify additional useful relevant information for systematically reviewing a trial programme?

We will primarily use descriptive methods to answer the questions. To answer question 1 we will compare the risk of bias in our 2012 review with risk identified after addition of M2 information to our current review using a 3 by 3 contingency table. We will repeat this procedure to answer question 2, by comparing risk of bias in our 2009 BMJ review to our current assessment. This analysis will be based on the subset of trials that were published and included in our 2009 review.

To answer question 3 we will list all the components of other risk of bias in the current review and compared these with previous reviews (2012 and 2009).

To answer the final question we will provide a summary of the items that were identified in our assessment of the trials using the new M2 tool. This will allow us to summarise discrepancies between what was planned in the protocol, what was carried out (RAP, protocol amendments), what was reported in M1, and what was reported in the published papers. The focus would be on the trial programme of research i.e. issues that appeared consistently over the trials.

Adverse events

2. Summary and background

This document outlines how we will conduct the analysis of adverse events as part of the wider Cochrane review of neuraminidase inhibitors (NIs) for prophylaxis and treatment of influenza in healthy adults and children (A159).

We use the term "adverse events" throughout this document rather than harms or adverse reactions as these latter terms imply causality which may or may not be appropriate.

In keeping with the methods of our previous review we will not use data from journal publications for this proposed analysis. We now have access to multiple clinical study reports (CSRs) for both oseltamivir and zanamivir. To our knowledge this is the first time some of these data have been available outside manufacturers and regulators, and allows for the exploration of events in more detail than is possible using the limited information on safety reported in journal publications. This potentially allows us to address some of the concerns that have arisen in the post marketing period about the possible relationship between neuraminidase inhibitors, oseltamivir in particular, and neuropsychiatric and other harms. The documents available to us contain listings and summaries of adverse events recorded in the trials including narrative summaries of serious adverse events and adverse events leading to study withdrawal.

The adverse events are classified by relationship to the study drug and also, by intensity (mild, moderate, severe, life-threatening, and death). The duration of events is reported and they are also lumped into body systems such as gastro-intestinal, neurological, etc.

2.1 Methods

All CSRs of oseltamivir and zanamivir will be included in our analysis. CSRs for prophylaxis, for treatment of adults, and for treatment of children will be analysed separately. Adverse events will be initially descriptively compared over the entire treatment and follow up period but then potentially stratified by on-treatment and off-treatment periods if it appears there may be a difference between treatment groups.

2.2 Adverse events for comparison

2.2.1 Common events

For common events of any intensity with an overall incidence of 2% or more we will compare the incidence between treatment groups. The cut-off of 2% is based on a power analysis where assuming 4000 patients in total (this is approximately how many patients we have access to in oseltamivir treatment trials of adults as well as in oseltamivir prophylaxis trials of adults), we will have 80% power to detect an odds ratio of 1.75 with 5% level of significance.

2.2.2 Uncommon events

Due to a lack of data to compare uncommon events we will compare events lumped into body systems between treatment groups. If we find evidence of a difference in incidences between groups lumped into a body system we will conduct further analysis if appropriate. This further analysis is to determine whether the difference in incidence is due to any common events included in that body system. For example in the case of neurological body system, if we found evidence of a difference between treatment groups we would remove all common neurological events such as headaches and repeat the analysis.

2.3 Severe, serious events and events leading to study withdrawal

As well as the analysis described in section 2.2 above we will also conduct a subgroup analysis of just the events with severe intensity, serious events and events leading to study withdrawal. We will use the same definitions of “severe” and “serious” as specified in the CSRs. However we will check the classifications using all the information available in the CSRs including line listings of events, narratives provided for serious events and also for events leading to study withdrawal. Any disagreements with the original classifications will be recorded and any reclassifications will be assessed in a sensitivity analysis. Given it is unlikely there will be sufficient events to conduct separate statistical analysis at the level of body system we will compare the overall distribution of events by body system between treatment groups.

2.4 Incidence of adverse events in the CSRs

As a further check on the validity of the data on adverse events contained in the CSRs we will conduct descriptive comparisons of the incidence of adverse events in the prophylaxis and treatment trials.

This is because of the unclear methods of collecting and classifying adverse events in the trials. A potential adverse event could have been classified as a symptom of influenza, an efficacy outcome (such as complication of influenza), or an adverse event. Hence an informal comparison of the incidence of adverse events in the trials where participants had influenza (or influenza-like-illness) and the trials where participants did not have influenza may help show where adverse events could have been under-reported. We will take into account factors such as age of participants and duration of treatment exposure for these informal analyses. In addition if it is clear that an adverse event was not reported as an adverse event but was included elsewhere in the CSR (e.g. in the efficacy section), we will include that data in our adverse event analyses.

We will also construct a table showing the definitions specified in each CSR for classifying potential adverse events as adverse events, complications or symptoms of influenza.

2.5 Antibody titre

We have already reported that antibody production was lower in the oseltamivir group than in the placebo group in the systematic review of treatment trials of oseltamivir (2012). We will update this analysis by including additional oseltamivir trials as well as assess antibody production in the zanamivir trials.

We will assess antibody production in the prophylaxis trials of oseltamivir and zanamivir by the following methods.

We will first identify the participants who had influenza-like illness (ILI) or pyrexia. If the proportion is similar between active group and placebo group, the proportion of participants who had four times or higher increase of antibody will be compared between groups.

2.6 Dose-response analysis

A number of trials included two or more active treatment arms with different doses of study medication given to participants in each of the arms. For these trials we will investigate the dose-response relationship for common adverse events (as defined above).

2.7. Details of analysis

Initial analysis will be descriptive only where we will report the numbers and percentages of events by treatment group. If there is a potential difference in the pooled percentages between treatment groups (e.g. if there is more than a two standard error difference between percentages) then we will conduct formal meta-analysis. If indicated we may also conduct additional analyses taking into account event intensity and/or duration.

2.8 Limitation and exploratory analysis

The methods presented above are those that we have pre-specified prior to formal analysis of the data. A limitation of these methods is that we may fail to detect differences in rare adverse events because these events will be compared along with other types of events within body systems. Therefore in the process of conducting our formal analysis we may generate further hypotheses or conduct additional exploratory analyses. If this is the case then we will clearly label these analyses as exploratory and interpret the findings accordingly.

Types of outcome measures

3. Background

For most people, influenza is a self-limiting illness. However the disease can at times lead to serious complications such as pneumonia and hospitalisations, and if treatment with neuraminidase inhibitors can reduce the risk of severe outcomes, this would be an important public health benefit. Another potentially important public health benefit would be the ability of antivirals to interrupt person to person transmission of influenza. Current evidence for these outcomes is scarce or inconclusive. A positive balance of effects on complications and viral spread versus harm profile is the main reason for using NIs in a public health context, especially the orally administered oseltamivir.

All analysis will be based on the intention-to-treat (ITT) or safety populations as our prior review discovered compelling evidence that the ITTI (the subpopulation deemed to be influenza infected) populations were not balanced between treatment groups in the Roche oseltamivir trials. In addition, estimates from the ITT population will be more generalisable to clinical practice where routine testing for influenza is not common in many countries (and even where used, remains of variable accuracy). Analysis will be conducted separately for prophylaxis trials, treatment trials of adults and treatment trials of children.

The list of outcomes given below includes all potential outcomes that we believe are clinically important. However a number of them may not be formally comparable in this review because there are insufficient numbers of events (e.g. mortality) or they were not adequately measured or reported (e.g. drug resistance).

3.1 Outcome measures for treatment studies

Complications~
Harms*
Symptom relief
Hospitalisation
Viral excretion
Drug resistance
Mortality

3.2 Outcome measures for prophylaxis studies

Influenza-like-illness^
Complications~
Harms*
Hospitalisation
Viral excretion
Drug resistance
Mortality

~Complications (secondary illnesses) include pneumonia, bronchitis, otitis media, sinusitis or other respiratory tract infection after influenza-like-illness. Initially we will construct a table to illustrate the design methodology used for each study. The table will include the following variables:
Study/trial ID
Where complications are first defined in the CSR (e.g. "as secondary endpoint in 3rd version of protocol six months into trial and two months prior to trial unblinding")
Definition of “complication” including types of events, population and time period at risk
How complications were measured (see diagnosis methods criteria shown below)
Availability of complications data for the ITT population

We will then stratify our analysis by method of diagnosis with three possible criteria:
a. Lab-confirmed diagnosis (e.g. based on radiological or microbiologically confirmed evidence of infection).
b. Clinical diagnosis without laboratory confirmation (diagnosed by a doctor after a clinical examination).
c. Other type of diagnosis such as self-reported by patient

*A separate section provides the details of our proposed analysis of harms.

^The main outcome of interest is any symptomatic influenza-like-illness (ILI). However, we will also conduct separate analyses of influenza (symptomatic and asymptomatic) and non-influenza ILI.

Contributors

Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ

COCHRANE COLLABORATION
ACUTE RESPIRATORY INFECTIONS REVIEW GROUP

Referee assessment of a Review, submitted for publication on *The Cochrane Library*

Title of Review:	Neuraminidase inhibitors for preventing and treating influenza in healthy adults
Review ID:	A047
Referee report due:	15 February 2006
Name of external referee:	[redacted]

Please use the accompanying checklist for guidance on the important areas to be covered in the review, and include any additional comments as required. Any specific suggestions that you can make for desired changes are greatly appreciated.

In the light of the assessments you make using the checklist, please give an overall opinion of the review (check one box only):

Acceptable for publication in its present form	☐
Acceptable for publication with minor revisions	☐
Acceptable for publication with substantial revisions	X
Not acceptable, major revisions needed	☐

COMMENTS FROM REFEREE

Thank you for asking me to referee this review. My comments are as follows:

Major comments:

Structure and presentation of planned analyses. At the moment multiple comparison are reported in the text and Forest plots. It makes it very hard for the reader to get an overall idea as to the structure of the review. I think it would be helpful if the authors maintain a consistent structure in the aims/objectives, outcome measures, and results section with regard to the multiple different outcomes they have assessed. The authors should state throughout the review- abstract, methods, results that studies were divided into prophylaxis, treatment and post-exposure prophylaxis and that these groups were further sub-divided according to case definition of “influenza” and “influenza-like illness”. Ordering presentation should be kept consistent throughout the review (at present this is not the case). They also present results for “influenza” further sub-divided into “symptomatic” and “asymptomatic” cases. Effect on “viral load” is also described. Lastly, different doses of drugs and formulations of drugs are used. The effect is to produce a well-described review at the expense of a review that is easy to follow and understand. Structuring the review with consistent ordering and format would be a help to the reader. See GLOSSARY, its introduction should aid comprehension.

The phrase “We structured the comparisons into prophylaxis, treatment and adverse events and further subdivided them by outcome and dose” has been added to the abstract and Data synthesis texts. The text wherever possible is now divided accordingly.

My other major concern relates to the fact that no synthesis of data for either agent has been performed. What is the justification for this? Surely NI’s as a drug class do not have differential efficacy (even though their mode of delivery differs)? I can understand why side effects may differ due to the different modes of delivery but am unsure why efficacy would be any different.
This has already been done (see aggregate estimate or bottom line of 01.01 as an example)

Linkage of reporting of results in the text with results in Forest plots- at present because of the multiple numbers of Forest plots, it is hard for the reader to follow where textual results are linked/derived from the Forest plots. We have never inserted links to Forest plots in any of the reviews – in addition the comparisons are clearly labelled “prophylaxis” and “treatment”. We are not sure what the editor/referee is getting at here.

Prevention of complications should be added as a key objective to the review and this should also be included in the text of the background section. “and their complications” has been added to the objectives.

Discussion- the discussion does not summarise the findings of the review; place them in context of other studies or discuss the shortcomings of the review. Whilst the authors concentrate on discussing NIs in seasonal and avian outbreaks (understandable in the context of what is likely to be considerable interest in this review), I feel that their points are hard to follow as the “bottom line” about prophylaxis, treatment and post-exposure prophylaxis have not bee summarised and the robustness of their findings not critically discussed.
We are not sure we agree with this comment – see the first line of the Discussion “We have assembled a good-quality uptodate evidence base of the prophylactic and treatment effects of NIs. These compounds have low effectivness, high efficacy and appear well tolerated, with the possible exception of oseltamivir-induced nausea and vomiting and zanamivir-induced diarrhoea” etc etc. We do not think that repeating our numerical results again (if this is what the referee is suggesting) in this section would enhance readability. Perhaps the addition of the GLOSSARY now makes things clearer – see what you think.

Discussion- seasonal influenza. The second sentence of the discussion states that NIs “have low effectiveness and high efficacy”. The authors need to clarify what the definitions of efficacy and effectiveness are in their review see GLOSSARY. Their subsequent arguments about use of NIs in serious epidemics/pandemics follows on from these definitions and are is hard to follow. The authors use the same terms in the “outcomes” section of the Table included studies”. Again I am unsure of the definition of the terms in this context as well. see GLOSSARY and our comments in response to consumer referee.

If the structure of the Cochrane Database of Systematic Reviews was clearly thought out we would have overviews i.e. for whole topics like influenza we would have an epidemiological introduction and an algorithm to aid decision making. With the current structure it is impossible to introduce the epidemiology of a topic (i.e. clarify the epidemiology of influenza and ILI) without considerably lengthening the review. We would then have the opposite criticism of too long, too complicated.

We have added the following which we hope will clarify our conclusions: “Existing trials on NIs were clearly designed and undertaken within a registration and regulation perspective. This is reflected in the cryptic reporting of continuous outcome data which forced us to resort to summary measures such as HR, which although methodologically virtuous, may not be relevant to workers in the field. NIs affect influenza symptoms, either preventing their appearance or

curtailing their duration and although we found clear evidence of their evidence in the interruption of transmission of seasonal influenza in households, NIs do not prevent infection and decrease but do not interrupt nasal shedding of seasonal influenza viruses. We cannot explain how NIs can affect respiratory complications of seasonal influenza such as bronchitis and pneumonia while not preventing infection and this effect should be further studied.

Discussion- role in avian influenza. The authors discuss the role of NIs in avian influenza. My concern is that they are introducing non randomised evidence in the context of this review and have (understandably) discussed efficacy of HIs in this context. Could this be going beyond the scope of the aims and objectives of this review. Most Cochrane reviews that I have reviewed or have been involved with have been very cautious and limited in their comments concerning the generalisability of the reviews' findings. We would have gone beyond the scope of the review had we introduced evidence from case-series into the review. We did this exactly for the reasons the referee gives in one of his earlier comments: to put the evidence in context.

In addition we now clearly state that evidence from seasonal influenza may not be generalisable to pandemic influenza or avian influenza "and we have doubts as to the generalisability of the evidence from seasonal influenza to avian influenza".

Minor comments:

Throughout the text some references have missing brackets around them. The text of the review needs to be proof-read to amend this error and make citation/presentation of references consistent throughout the review. Liz – can you help please. We did not insert brackets when the reference was either the subject or the object of the phrase.

Synopsis- there is no full stop after the first sentence. There is a gap after the first sentence and the following paragraphs. I think the authors are making several different points in their last sentence that are not clearly related; it would be clearer if this sentence was broken into separate statements. This has been done.

Abstract- results, it would be helpful if the number of included RCTs are stated at the beginning. We have inserted: "We identified 4 prophylaxis trials, 13 treatment trials and two PEP trials" I think it would be helpful if the prior subgroups of analysis were stated in the "data collection and analysis", clearly outlining the subheadings under which the results are reported (see above). This has been done (see earlier response).
Whilst I do not disagree with their last statement in the "reviewers conclusions", I think it is speculative and not based on any data in the review. Fair enough but that is our view, and we are allowed to express it.

Background- final sentence of para 2, prevention of complications should be added. Done

Objectives- prevention of complications should be added as an objective. It was already in the abstract, and we have inserted it in the main text as per earlier comment.

Types of participants- why is the criterion to include studies only if "no less than 75% of the subjects are aged 14 to 60" used? I presume that this is to exclude more elderly subjects who are at higher risk of complications. This phrase has been inserted It would be helpful if the authors clarified this point and outline the number of studies in older patients that were excluded on this basis. See response to stats referee.

Outcomes measures- prevention of complications should be added as an outcome measure. Done

Methodological quality of included studies- I do not understand what the authors mean in paragraph 3 of this section. I cannot find the outcomes they have numbered in this paragraph amongst the Forest plots. We have taken the sentence out.

Dose response relationship- there doesn't appear to be a dose-response relationship with Oseltamivir. We could not find the relevant text.

Referee's name: [redacted]

Date: 26 Jan 2006

TO: [redacted]

FROM: [redacted]

DATE: April 10, 2006

RE: Comments from referee

Main Results

The statement that NIs actually increase ILIs is surprising since multiple studies show consistent evidence of 70-90% efficacy for household members and for seasonal flu in the community (NEJM 2005;353:1367 Table 3 with 6 studies). Am I misunderstanding the analysis? (See below) We do not understand this comment

Results – prophylaxis

Does it make sense that NIs increase the rate of ILIs (RR 1.3) and decrease the rate of symptomatic flu (RR 0.4). It is not only the paradox of a difference, but the magnitude of the difference. Also, if it is given for flu, isn't it fair to expect a change only in flu? The 60% protection rate is the number always quoted and this is the basis for stockpiling for Avian flu. Our version of the review reports "Compared to placebo, NIs have no effect against ILI (RR 1.28, 95% CI 0.45 to 3.66 for oral oseltamivir 75 mg daily, RR 1.51, 95% CI 0.77 to 2.95 for inhaled zanamivir 10 mg daily)", i.e. NIs do not work against ILI, nowhere do we state that they increase ILI rates.

Discussion

The major side effect of zanamivir is cough and bronchospasm. May be but not from the trials included in the review.

The Kiso study needs to be interpreted in the context that it was done in children who have high rates of resistance and in Japan which has the highest usage rates of oseltamivir in the world. "a country with very high NI prescription rates" has been inserted in the text.

It does not make sense to review the data from the outbreak of H7N3 in Canada since "effects of osteltamvir were outside a formal study". This the Discussion, not the results section of the review – see comments and responses to Tom Fahey.

It is okay to say use of these drugs in pandemic flu is not supported by any credible data. However, it should also say that studies to show efficacy for treatment or prevention of pandemic flu have not been done. We have added "Further research on the possible effects of NIs on avian influenza subtypes is also required" to the implications sections.

Conclusions

The two sentence conclusion seems too simplistic. Most countries that can afford NIs are stockpiling with the assumption that they may work. There is nothing in the review to say they are wrong. This point is not made well. No, we are saying that they should not be used on their own in a pandemic, not that they should be used at all.

We have inserted the underlined rider: "Finally, the inability of the NIs to prevent infection and to suppress viral nasal excretion raise doubts as to their effectiveness in interrupting viral spread in a pandemic, although NIs may have a role in addressing symptoms and complications. We conclude that in a pandemic, NIs should be used within a package of measures to interrupt spread, that is to say, together with barrier, distance and personal hygiene measures.

There should be more attention to the potentially important role for these drugs in reducing viral load to reduce transmission. This is the rationale for use in hospital admissions – proven to reduce viral shedding and not proven to reduce transmission but very logical. Yes, but we have presented all the evidence we could find. They would appear to reduce transmission in families for seasonal influenza – in a pandemic it's anybody's guess.

Title of the review:

A047 Neuraminidase inhibitors for preventing and treating influenza in healthy adults

Authors:

Jefferson TO, Demicheli V, Di Pietrantonj C, Jones M, Rivetti D

This is an important review especially in the context of present epidemic. What I understand is no more studies have been added in this review except changing some terminologies. The authors should be complemented on the detailed analysis and extensive statistical analysis. Except first three tables, all other tables represent single studies. Even though quasi randomised trials are expected to be included in the review, I do not find any such study in the review. This will further enhance validity of the findings.

Types of participants:
"Individuals with no known pre-existing chronic pathology known to aggravate the course of influenza in studies in which no less than 75% of the subjects are aged 14 to 60."
This is slightly confusing (at least for me). Why this criteria and how many such studies excluded due to not satisfying the above criteria?

None – see table of excluded studies

The text now reads "In keeping with our objective of reviewing evidence on healthy adults, we only considered studies in which no less than 75% of the subjects are aged 14 to 60".

Otherwise all statistical points are clear.

Checklist for refereeing a review

The aim of the checklist is to assist an editor/external referee in identifying any areas of weakness in a review, and suggesting improvements. This checklist is provided as a guide only. Please make any comments about the refereed review on the accompanying assessment form.

1) BACKGROUND

The background section is designed to explain to people what is going to be reviewed and why. It must explain why the question needing to be answered is important. For example, it should indicate the areas of uncertainty in relation to the intervention and highlight issues that are controversial or the subject of public concern. It must define all terms and interventions clearly, and should try to use a balanced tone that does not pre-judge the value of the intervention.

The background should be brief. As a guide, it should be a page long. It is not a monograph or an overview and should be concise and clear.

Has the Author covered these issues? ☐ **Yes** ☒ **Partially** ☐ **No**

2) OBJECTIVES

The objective(s) should be clear and specific with a precise statement including the intervention(s) reviewed and the targeted problem. The objectives should be specific and consider the three elements of:

- population / types of participants
- types of interventions / comparisons
- types of outcome measures of interest

Has the Author covered these issues adequately? ☐ **Yes** ☒ **Partially** ☐ **No**

3) CRITERIA FOR CONSIDERING TRIALS FOR THE REVIEW

The criteria used to select trials for inclusion in the review should be stated. They should specify:

- types of studies
- types of participants
- types of intervention(s)
- types of outcome measure

Has the Author covered types of study design? ☒ **Yes** ☐ **No**

Has the Author covered types of participants?	[x] **Yes**	[ ] **No**

But exclusion of RCTs of elderly patients should be explained and quantified (see above).

Has the Author covered types of intervention(s)?	[x] **Yes**	[ ] **No**
Has the Author covered types of outcome measure?	[x] **Yes**	[ ] **No**
Are any reasons for excluding studies clearly reported?	[ ] **Yes**	[x] **No**

4) THE RESULTS

Were the methods used (whether qualitative or quantitative) to combine the findings of the relevant studies clearly reported?	[x] **Yes**	[ ] **Partially**	[ ] **No**
Were the methods used to combine findings appropriate to the questions in this review?	[x] **Yes**	[ ] **Partially**	[ ] **No**
Are results sensitive to changes in the way the analysis was done?	[ ] **Yes**	[x] **Partially**	[ ] **No**
Is the precision of results reported relevant?	[x] **Yes**	[ ] **Partially**	[ ] **No**
Were data limitations and inconsistencies discussed adequately?	[ ] **Yes**	[ ] **Partially**	[x] **No**
Was the summary of the findings adequate?	[ ] **Yes**	[ ] **Partially**	[x] **No**

5) CONCLUSIONS

Are the conclusions clear?	[ ] **Yes**	[ ] **Partially**	[x] **No**
Are the conclusions supported by data and/or analysis in the review?	[ ] **Yes**	[x] **Partially**	[ ] **No**
Are important considerations for decision makers identified?	[ ] **Yes**	[x] **Partially**	[ ] **No**

6) IMPLICATIONS FOR PRACTICE

If there are implications stated are they justified?	☐ **Yes**	☒ **Partially**	☐ **No**
Are the most vital issues included?	☐ **Yes**	☒ **Partially**	☐ **No**

7) IMPLICATIONS FOR RESEARCH

If there are implications stated are they justified?	☐ **Yes**	☒ **Partially**	☐ **No**
Are the most vital issues included?	☐ **Yes**	☒ **Partially**	☐ **No**

If NO, what other issues should have been included?

8) THE ABSTRACT

The abstract is the most frequently read part of the review and will be translated into many languages. It will also be available on MEDLINE and other databases.

Are the most important findings and conclusions included in the abstract?	☐ **Yes**	☒ **Partially**	☐ **No**
Is the abstract consistent with the full review?	☒ **Yes**	☐ **Partially**	☐ **No**
Is the style of writing easy to understand?	☐ **Yes**	☐ **Partially**	☒ **No**
Is the style of writing interesting	☐ **Yes**	☒ **Partially**	☐ **No**

9) PLAIN LANGUAGE SUMMARY (SYNOPSIS)

The plain language summary (formerly called the 'synopsis') aims to summarise the review in an easily understood style which would be understandable by consumers of healthcare. The first part is a restatement of the review's title using plain language terms.

The second should be no more than 400 words in length and should include:

· A statement about why the review is important: for example definition of and background to the health care problem, signs and symptoms, prevalence, description of the intervention and the rationale for its use.

· The main findings of the review: this could include numerical summaries when the review has reported results in numerical form, but these should be given in general and easily understood forms. Results in the plain language summary should not be presented any differently from in the review (i.e. no new results should appear in the summary). Where possible an indication of the number of trials and participants on which the findings are based should be stated.

· A comment on any adverse effects.

· A brief comment on any limitations of the review (for example trials in very specific populations or poor methods of included trials).

Has the Author covered these issues? ☐ **Yes** ☒ **Partially** ☐ **No**

10) CONFLICT OF INTEREST:

Do you have any concerns about possible conflicts of interest in this review? ☐ **Yes** ☐ **Partially** ☒ **No**

If YES or PARTIALLY, what concerns do you have?

10/80/01 – Anonymised Peer Review Comments

Reviewer 1

This application seeks additional funding to complete a systematic review which has grown in size since the previous funding award - owing to the manufacturer releasing further trials and data. Given the complication of serial awards and the number of options outlined in the previous application (I am uncertain exactly which options and components were commissioned), it is difficult to tease out what is work outstanding and what is new work associated with having received more trials/information than anticipated.

The project has two concurrent main aims: (i) systematic review and meta-analysis of neurimidase inhibitors and (ii) an evaluation of the integrity of conduct and reporting in the Roche funded/ sponsored trials of oseltamavir. It is unclear if the applicants will subject the materials provided by GSK on zanamavir to the same level of scrutiny - to be even handed they should be.

The methods for the systematic review outlined in the protocol appear appropriate. The applicants decision to base analyses solely on the clinical study reports is sensible given the extent of unpublished material (this is stated only for adverse events but I assume that it will be the same for all outcomes), as is the decision to focus on the ITT group meaning that findings are more likely to be generalisable to the population at large.

For the analyses of antibody titre (which I believe was initiated on the basis of observations during the previous phase of the project i.e. was not pre-specified), the applicants plan to compare treatment arms using the proportion of patients with a four times increase in antibody level. It is unclear how the particular (4x) cut-point was chosen, whether this value has particular biological significance and/or was selected independent of data analysis. This is particularly important because as these analyses were not planned at the outset of the project and appear to be somewhat data driven, they are open to criticism.

Most effort to date appears to have been in receiving and organising documentation and information received from Roche. The applicants have received a vast amount of information. However, the number of pages is not necessarily indicative of the effort required for data extraction. It is likely that the clinical study reports contain tabulated data that are amenable to the intended clinical analyses.

Much of the proposed activity is associated with data extraction and cross checking across data sources, comparing consistency across company documents as well as with published reports. It appears that the applicants plan to cross check the categorisation and collation of adverse event against the patient narratives contained in the clinical study reports. This is a very large undertaking given that as many as 4000 patient narratives may be available and that some clinical input may be required. I suspect that to complete the task within the requested additional time and resources, some sort of compromise may be necessary. For example, checking whether categorised and collated outcomes accurately reflect the patient narratives contained in the clinical study reports and case report forms, could be carried out for a random sample of individuals rather than for all 4000 individuals included in the trials. A stepped approach might be helpful. If investigation of the narratives in a sample of individuals reveals only minor inconsistency, then the applicants might be able to proceed directly to analyses based on CSR tabulated data.

If verification of all patient narratives is required, then the applicants would be able to reconstruct the trial individual participant-level data with respect to key baseline characteristics and outcomes

as part of their extraction and verification processes. If so, then a valuable addition to the work would be to carry out subgroup analyses of key types of patients. This could be particularly useful for at risk populations such as the very elderly or those with asthma or chronic bronchitis.

It is unclear how informative some of the detailed scrutiny of documentation will be, for example the grid outlining development of the protocol and analysis plan. I suspect that much in the M2 documents will be irrelevant. Similarly it is unclear what the tools under development are and whether they will be of wider usefulness. I was unable to work out whether the tools are process descriptors or some form of automated data extraction tool. This aspect of the work is arguably tangential to the main project, but may be necessary to complete the project - with more general application a potential spin-off.

On the whole, this is an important project that has the potential to be very high profile. It is currently part-done and stopping with only a likely biased sample of the trials analysed would be wasteful of the resources already invested. I recommend that the additional funding is granted. However, I have some concerns that the applicants will not be able to achieve the level of scrutiny that they plan within the additional time and resource requested. As I imagine that the clinical results are of most importance to the HTA programme, it may be sensible to request that the applicants prioritise completion of the analyses of clinical outcomes of importance to patients (as listed in their protocol) based on all randomised trials over some of the detailed investigation of reporting across company documents. It may also be useful to confirm timelines around receipt of outstanding trials from the manufacturer to ensure that these align with the timelines and to ensure completion without recourse to a further extension.

Reviewer 2

I think the 12-page request for additional funding is poorly constructed and does not provide a well-argued case. In the time available, I regret I have not been able to scrutinise all of the relevant documents in as much detail as I would have liked to have done. There seem to be three potential advantages of doing the extra work:

1. Changes to the substantive treatment estimates of the review from the addition of new data not previously available, e.g. because data from commercial trials had been withheld. This advantage should probably be split into two (1a and 1b), i.e. benefits and harms/adverse effects.
2. Changes to the INTERPRETATION of the treatment estimates of the review arising from the influence of better information about the conduct of trials on the assessment of risk of bias.
3. The development and assessment of a tool to extract and appraise information from similar study reports in the future, for different topics.

1a Beneficial/intended treatment effects

It's not clear to me whether the additional documents just received represent more / better data that would contribute to aggregate treatment effects for the intended outcomes, i.e. symptom scores, risk of 'flu, transmission risk, etc. If yes, this would be a major advantage, reducing the risk of biased aggregate treatment effects arising because of selectively withheld non-significant trial results.

1b Harmful/unintended treatment effects (adverse effects)

I think it is clear that the existing evidence about possible harms of the drugs does not include the information contained in the additional study reports. Therefore, if adverse effects of these drugs are suspected, this output would be a very important one. The approach described by the applicants

seems fairly standard and I don't see any reason why this output would not be delivered.

I presume that the "body system" classification is MedDRA or a similar industry-standard coding system. The applicants need to be aware that there may be some quirks of the coding system; for example, pulmonary embolism is classified as respiratory, not cardiovascular. This may hinder comparisons by organ system if suspicion is raised for a particular SAE which is 'oddly' classified (since a differentially elevated frequency may be diluted among the noise of other unrelated SAEs). (This point is made / acknowledged under 2.8.)

The pre-specified analyses (sections 2.2 to 2.7) are not specified in detail. There are conditional statements (e.g. if X then Y) based on "e.g." criteria (>2SE difference). If all of the relevant data have been extracted and entered, then the work involved in doing the meta-analysis is not very much. I would prefer decisions about the meta-analysis to depend on the appropriateness of pooling / importance of the outcome rather than on the potential to demonstrate a statistically significant difference.

2 Changes to the interpretation of the treatment estimates

I would expect the additional, detailed information in the study reports to clarify aspects of the original risk of bias assessment. The value of such clarification is difficult to anticipate in advance but may not be so great.

Referring to Figure 1 in the existing review, the main areas for clarification appear to be in relation to random sequence generation, other bias and (to a lesser extent) blinding of outcome assessment. I would imagine there is a high prior probability of random sequence generation in trials designed and sponsored by industry, so I wouldn't expect the additional information to add much. I am not sure what 'other biases' are of concern. Most trials appear to have been blinded so (without spending a lot of time studying the review in detail) I am confused why there may have been doubt about blinding of outcome assessment.

Surprisingly, the risk of selective reporting (potentially a serious risk since I believe it has been alleged that the manufacturers selectively withheld trial results) appears to have been clear (either high or low risk) in most of the included trials.

3 The development and assessment of a tool

I am sure that this is do-able (the team are part of an experienced Cochrane review group) but I am uncertain about the value it will add. I don't get a feel for the methodological issues that would be informed – and that would generalise – to additional future topics from this careful scrutiny of a set of commercial trials. ("The focus would be on the trial programme of research i.e. issues that appeared consistently over the trials." pp.4-5)

Estimated amount of work / funding requested

It appears as though the applicants intend to extract the required data by hand. This would indeed be a Herculean task if the intention was to do some sort of individual patient data analysis. However, I don't think this is the case – and I suspect that the number of pages in the reports is not a particularly good guide to the work involved to address the questions above. The question about adverse effect looks most important to me. I think that the work required to extract these data from line listings should not be so great. Therefore, I feel that the amount requested is quite high, although I appreciate that the budget is intended to cover all three questions and that about half of the total is made up of indirect costs.

Summary
The aims are worthy but I think not all of them are central to the aims of the original project. The supporting document is not very clear about the current gaps in the review and the ways in which the additional data will help to resolve them. In my view the main output from the additional study reports will be clarification about the possible harms of the two drugs. This is very important – but, alone, may not require the level of resources/funding requested.

18/07/2013

Correspondence:

Prof Carl Heneghan
Dept of Primary Care Health Sciences
University of Oxford

Email: carl.heneghan@phc.ox.ac.uk

NIHR Health Technology Assessment Programme Project: 10/80/01 - Update and amalgamation of two Cochrane Reviews: neuraminidase inhibitors for preventing and treating influenza in healthy adults and children

Response to review comments:

Before any additional funds are released we require confirmation from you in writing that you have all of the data in your possession, are not awaiting further data, and that the data in your possession is of sufficient quality to complete your planned analyses.

Point 1: We are concerned that the new data supplied by Roche and GSK may contain patient identifiable material. We therefore request that you seek the opinion of an ethics committee and either supply evidence of ethical approval or confirmation that approval is not required.

1. Response:

The data we have received is compatible with data that has been made available to us from the European Medicines agency under their Freedom of Information policies. In addition, both GSK and Roche have confirmed their strong commitment to protecting patient identity and that the data they are releasing to us has been redacted and is happening unconditionally and we are free to do what we want with the data as it does not contain patient identifiable material.

In de-identifying data there are two methods. One is the appropriate use of the Harbor method (the safer method) which includes:

Taken from Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html#preparation)

C) All elements of dates (except year) for dates that are directly related to an individual, including birth date, admission date, discharge date, death date, and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older

And

(R) Any other unique identifying number, characteristic, or code, except as permitted by paragraph (c) of this section [Paragraph (c) is presented below in the section “Re-identification”]; and

This is the methodology applied to the initial GSK and subsequent Roche CSR documents we have received.

However, the second method that can be employed requires expert determination:

'(b) Implementation specifications: requirements for de-identification of protected health information. A covered entity may determine that health information is not individually identifiable health information only if:

(1) A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable:

(i) Applying such principles and methods, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information; and

(ii) Documents the methods and results of the analysis that justify such determination;'

By recoding the serious adverse event narratives, we have come to a blended approach.

There are no issues that require ethical approval in such an approach and we are planning to publish the methods of this approach.

In addition, although we do not need to seek ethical approval as such for the data we have currently received, we will seek an ethical opinion to better understand the concerns that may hinder future projects using this type of approach.

Point 2: Your draft protocol indicates that you don't have the electronic data files and therefore must manually extract the data. Did GSK and Roche refuse to supply electronic data files? And if electronic data was available why did you choose to manually extract the data from PDFs?

2. Response:
We have received all the GSK data (30 CSRs).

We have received two tranches of the Roche data (updated table of contents is provided and appended to the additional funding protocol) and are continuing to receive these large files for processing. In our email correspondence we should have all of the phase 3 trials by the end of July, which we can confirm upon receipt by email.

Our present analysis does not focus on electronic data for several reasons:

First, Roche has only promised to make IPD available from Jan 1, 2014.
(See http://roche-trials.com/dataSharingPolicy.action)

Second, we are carrying out an analysis according to a protocol we have published quite some time back and that was written at a time when all Roche had promised (but not delivered in full) were "full study reports" for the ten trials that make up the Kaiser 2003 pooled analysis. Our expectation was that we would have full CSRs for those trials—and we hoped to obtain full CSRs for other trials.

Third, we asked Roche if it would consider entering into a contract similar to the one they entered into with Harvard researchers (who did receive electronic patient-level data), and Roche did not respond to this query.

Fourth, even having IPD, it is not clear that it would result in an analysis superior to the one we're doing; and it seems very clear that for the outcomes of particular public health relevance--complications, hospitalizations--it does not appear necessary at all given the nature of these outcomes in the trials (few events and definitions were loose) so manual extraction from PDF does not require enormous effort. Roche's comments in reported in the BMJ would seem to confirm this (http://blogs.bmj.com/bmj/2013/06/26/trish-groves-what-does-tamiflu-do-and-how-will-we-know/).

Fifth, a full IPD meta-analysis of symptom reduction, for example, would be a major undertaking and would require a larger budget than we have at this point. We are not convinced such data are required to establish the overall safety and effectiveness of the named interventions. There may be instances where such analyses may prove useful, such as time to event data, but they are currently not warranted.

Point 3: Is the development of a data extraction tool required in order to complete the work or is this a piece of extra work? If this is a piece of extra work what proportion of the costs requested will be used to develop this tool?

3. Response:
These tools have already been developed and tested, and are attached as appendices.

Point 4.1 : Peer review suggested that the protocol document could be significantly improved. In particular it is not sufficiently detailed and comprehensive in describing and guiding your work. Two sets of peer review comments have been anonymised and attached to this letter. Please supply an amended version of your protocol taking these comments into account.

4.1 Response:
We have supplied an amended version of the protocol for further funding. The full protocol is still available for scrutiny on the Cochrane library. .

Point 4.2: The project has two concurrent main aims: (i) systematic review and meta-analysis of neurimidase inhibitors and (ii) an evaluation of the integrity of conduct and reporting in the Roche funded/ sponsored trials of oseltamavir. It is unclear if the applicants will subject the materials provided by GSK on zanamavir to the same level of scrutiny - to be even handed they should be.

4.2 Response:
Both sets of data will be, as suggested, dealt with in an 'even handed' way. Indeed this one of the main reasons for the extra additional funds, as in our original funded project we were not in receipt of all of the GSK data (it had not been submitted to the EMA for approval, therefore they had no holdings of zanamivir data), and until recently were not in receipt of the full set of clinical study reports.

In terms of the integrity our aim is to assess the methodological quality of the studies as well as determine the overall risk benefit profile.

Point 4.3:
The methods for the systematic review outlined in the protocol appear appropriate. The applicants decision to base analyses solely on the clinical study reports is sensible given the extent of unpublished material (this is stated only for adverse events but I assume that it will be the same for all outcomes), as is the decision to focus on the ITT group meaning that findings are more likely to be generalisable to the population at large.

4.3 Response:
This is correct and no response needed.

Point 4.4:
For the analyses of antibody titre (which I believe was initiated on the basis of observations during the previous phase of the project i.e. was not pre-specified), the applicants plan to compare treatment arms using the proportion of patients with a four times increase in antibody level. It is unclear how the particular (4x) cut-point was chosen, whether this value has particular biological significance and/or was selected independent of data analysis. This is particularly important because as these analyses were not planned at the outset of the project and appear to be somewhat data driven, they are open to criticism.

4.4 Response:
This is the standard diagnostic technique for determining influenza. It has been used by the manufacturers and also by NICE in their decision making, and is standard practice for the WHO:

Serological detection of influenza (WHO laboratory procedures)

'Serological methods rarely yield an early diagnosis of acute influenza virus infection. However, the demonstration of a significant increase in antibody titers (greater than or equal to 4 fold) between acute phase and convalescent phase sera may establish the diagnosis of a recent influenza infection even when attempts to detect the virus are negative. Apart from their retrospective diagnostic value, serological methods such as virus neutralization and haemagglutination inhibition are the fundamental tools in epidemiological and immunological studies, as well as in the evaluation of vaccine immunogenicity.'

Point 4.5: Most effort to date appears to have been in receiving and organising documentation and information received from Roche. The applicants have received a vast amount of information. However, the number of pages is not necessarily indicative of the effort required for data extraction. It is likely that the clinical study reports contain tabulated data that are amenable to the intended clinical analyses.

Much of the proposed activity is associated with data extraction and cross checking across data sources, comparing consistency across company documents as well as with published reports. It appears that the applicants plan to cross check the categorisation and collation of adverse event against the patient narratives contained in the clinical study reports. This is a very large undertaking given that as many as 4000 patient narratives may be available and that some clinical input may be required. I suspect that to complete the task within the requested additional time and resources, some sort of compromise may be necessary. For example, checking whether categorised and collated outcomes accurately reflect the patient narratives contained in the clinical study reports and case report forms, could be carried out for a random sample of individuals rather than for all 4000 individuals included in the trials. A stepped approach might be helpful.

If investigation of the narratives in a sample of individuals reveals only minor inconsistency, then the applicants might be able to proceed directly to analyses based on CSR tabulated data.

If verification of all patient narratives is required, then the applicants would be able to reconstruct the trial individual participant-level data with respect to key baseline characteristics and outcomes as part of their extraction and verification processes. If so, then a valuable addition to the work would be to carry out subgroup analyses of key types of patients. This could be particularly useful for at risk populations such as the very elderly or those with asthma or chronic bronchitis.

Point 4.5 Response:

The reviewer has overestimated the amount of work involved in our review of patient narratives. There are approximately 13,000 patients across the oseltamivir treatment and prophylaxis studies. Patient narratives are only available for those patients that had serious adverse events or withdrew from the study early.

Part of our extraction includes capturing the methods used to compare groups for primary and secondary outcomes and methods for additional analyses, such as subgroup analyses and adjusted analyses. As well as Ancillary analyses: results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory.

Therefore we have the ability to perform subgroup analysis and whilst they are important, it is important to not lose sight of the fact that the use and governmental stockpiling of oseltamivir is for its routine use in asymptomatic and symptomatic members of the community. Our review thus considers the evidence base that applies to the vast majority of people.

In addition, the power to detect important clinical outcomes, such as a difference in all-cause hospitalization is very small hence a subgroup analysis on this outcome would raise the risk of a type I error.

Point 4.6:
It is unclear how informative some of the detailed scrutiny of documentation will be, for example the grid outlining development of the protocol and analysis plan. I suspect that much in the M2 documents will be irrelevant. Similarly it is unclear what the tools under development are and whether they will be of wider usefulness. I was unable to work out whether the tools are process descriptors or some form of automated data extraction tool. This aspect of the work is arguably tangential to the main project, but may be necessary to complete the project - with more general application a potential spin-off. On the whole, this is an important project that has the potential to be very high profile. It is currently part-done and stopping with only a likely biased sample of the trials analysed would be wasteful of the resources already invested. I recommend that the additional funding is granted. However, I have some concerns that the applicants will not be able to achieve the level of scrutiny that they plan within the additional time and resource requested.

4.6 Response:
This raises an important question: is M2 irrelevant? Without actually performing this analysis we cannot answer such a question. However, our preliminary work on Module 2s, provided by the EMA, suggests that they are highly relevant. M2 contains the trial protocol and amendments for example. Nonetheless our work should provide an empirical answer to this important methodological question, but the reviewer is correct in that an assessment of the resource implications is important to understanding whether this level of scrutiny can be replicated in the future.

Point 4.7:
As I imagine that the clinical results are of most importance to the HTA programme, it may be sensible to request that the applicants prioritise completion of the analyses of clinical outcomes of importance to patients (as listed in their protocol) based on all randomised trials over some of the detailed investigation of reporting across company documents. It may also be useful to confirm timelines around receipt of outstanding trials from the manufacturer to ensure that these align with the timelines and to ensure completion without recourse to a further extension.

4.7 Response:
We are happy to do this.

Point 4.8:

I think the 12-page request for additional funding is poorly constructed and does not provide a well-argued case. In the time available, I regret I have not been able to scrutinise all of the relevant documents in as much detail as I would have liked to have done.

4.8 Response:
Our protocol stated at the outset:
Note: this document must be read in conjunction with the current version of A159 (Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database of Systematic Reviews 2012, Issue 1. Art. No.: CD008965. DOI: 10.1002/14651858.CD008965.pub3).
The overall methods and history of the review are reported in A159.

Without scrutinization of all of the relevant documents this may have led to some confusion. We have therefore provided one full protocol document for clarity. However, given the details outlined in A159 we would still emphasize that this document be read in conjunction with A159.

Point 4.9:
There seem to be three potential advantages of doing the extra work:
1. Changes to the substantive treatment estimates of the review from the addition of new data not previously available, e.g. because data from commercial trials had been withheld. This advantage should probably be split into two (1a and 1b), i.e. benefits and harms/adverse effects.
2. Changes to the INTERPRETATION of the treatment estimates of the review arising from the influence of better information about the conduct of trials on the assessment of risk of bias.
3. The development and assessment of a tool to extract and appraise information from similar study reports in the future, for different topics.

4.9 Response:
We agree with this assessment.

Point 4.10:
Beneficial/intended treatment effects
It's not clear to me whether the additional documents just received represent more / better data that would contribute to aggregate treatment effects for the intended outcomes, i.e. symptom scores, risk of 'flu, transmission risk, etc. If yes, this would be a major advantage, reducing the risk of biased aggregate treatment effects arising because of selectively withheld non-significant trial results.

4.10 Response:
It is becoming increasingly clear that data taken from journals, with strict word limits (often less than 3000 words) and often with restriction on tables, is inadequate in terms of assessing the intended outcomes and has introduced the kind of biases the referee alludes to.

Point 4.11:
Harmful/unintended treatment effects (adverse effects)
I think it is clear that the existing evidence about possible harms of the drugs does not include the information contained in the additional study reports. Therefore, if adverse effects of these drugs are suspected, this output would be a very important one. The approach described by the applicants seems fairly standard and I don't see any reason why this output would not be delivered.

4.11 Response:
We agree with this point.

Point 4.12:
I presume that the "body system" classification is MedDRA or a similar industry-standard coding system. The applicants need to be aware that there may be some quirks of the coding system; for example, pulmonary embolism is classified as respiratory, not cardiovascular. This may hinder comparisons by organ system if suspicion is raised for a particular SAE which is 'oddly' classified (since a differentially elevated frequency may be diluted among the noise of other unrelated SAEs). (This point is made / acknowledged under 2.8.)

4.12 Response:
We are aware of such quirks in the MedDRA system and duly note the reviewers concern.

Point 4.13:
The pre-specified analyses (sections 2.2 to 2.7) are not specified in detail. There are conditional statements (e.g. if X then Y) based on "e.g." criteria (>2SE difference). If all of the relevant data have been extracted and entered, then the work involved in doing the meta-analysis is not very much. I would prefer decisions about the meta-analysis to depend on the appropriateness of pooling / importance of the outcome rather than on the potential to demonstrate a statistically significant difference.

Changes to the interpretation of the treatment estimates
I would expect the additional, detailed information in the study reports to clarify aspects of the original risk of bias assessment. The value of such clarification is difficult to anticipate in advance but may not be so great.

4.13 Response:
This point presumes some element of experience with the handling of clinical study reports and their contents, I draw your attention to points 4.14 and 4.15 which illustrate the importance of greater scrutiny.

We have amended the protocol to detail the analysis issues outlined.

Point 4.14:
Referring to Figure 1 in the existing review, the main areas for clarification appear to be in relation to random sequence generation, other bias and (to a lesser extent) blinding of outcome assessment. I would imagine there is a high prior probability of random sequence generation in trials designed and sponsored by industry, so I wouldn't expect the additional information to add much. I am not sure what 'other biases' are of concern. Most trials appear to have been blinded so (without spending a lot of time studying the review in detail) I am confused why there may have been doubt about blinding of outcome assessment. Surprisingly, the risk of selective reporting (potentially a serious risk since I believe it has been alleged that the manufacturers selectively withheld trial results) appears to have been clear (either high or low risk) in most of the included trials.

4.14 Response:
We disagree with the reviewer's assessment of the potential for other biases. As an example, of the potential additional biases, that may have never come to light before and can only be found through scrutiny of M2 documents:

WV15673 and WV15967 were designed as two separate trials and were only combined after a protocol amendment due to the very low rates of influenza infection: WV15673 study dates were Jan 5th 1998 to April 3rd 1998 and WV15967, Jan 12th 1998 to March 27th 1998 - This amendment occurred on the 8th of June 1998, after both studies had completed.

There are methodological issues over the combining of these two trials. The placebo control event rates were very different: WV15673D control event rate was 2.4% compared with 7.1% in WV 15697D – The core report states that a pooled analysis was performed to check for consistency but these effects were not reported. Given the considerable heterogeneity and that one trial had only 12 influenza cases methodologically they should only be combined using meta-analytic techniques

In addition after the trial had completed a change to protocol amendment included lowering the temperature for the outcome definition. The reason given for was the use of additional antipyretics for the groups. Yet, there are substantial imbalances reported in terms of medication use, particularly anti-inflammatory drugs, which will have substantial effects on clinical presentation and potential outcomes.

In terms of outcome reporting there is an Errata based on the fact that adverse events were identified which failed to appear in either the 'on' or the 'off' treatment period – 14 events in total were erroneously reported.

Point 4.15:
The development and assessment of a tool
I am sure that this is do-able (the team are part of an experienced Cochrane review group) but I am uncertain about the value it will add. I don't get a feel for the methodological issues that would be informed – and that would generalise – to additional future topics from this careful scrutiny of a set of commercial trials. ("The focus would be on the trial programme of research i.e. issues that appeared consistently over the trials." pp.4-5)

4.15 Response:
We thank the reviewer for his comments about the team, in terms of the value to be added, this can only be answered by completing the project.

Point 4.16:
The aims are worthy but I think not all of them are central to the aims of the original project. The supporting document is not very clear about the current gaps in the review and the ways in which the additional data will help to resolve them. In my view the main output from the additional study reports will be clarification about the possible harms of the two drugs. This is very important – but, alone, may not require the level of resources/funding requested.

4.16 Response:
In terms of harms we draw attention to the recent research outlined by Rogers et al. 'Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion.'[1] The finding that 'information from all journal publications and conference abstracts could not identify a complete set of outcome data for any study' and 'only 19% of adverse events in the individual participant data have been reported somewhere in the published literature,' implies that systematic reviews based on journal publications alone are likely to be seriously flawed, in terms of assessing the trade-offs between the risks and benefits of the intervention.

[1] Rodgers MA, Brown JV, Heirs MK, Higgins JP, Mannion RJ, Simmonds MC, Stewart LA. Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion. BMJ. 2013 Jun 20;346:f3981. doi: 10.1136/bmj.f3981.

5. Reviewers also questioned your choice of a 4 fold rise in antibody titre as being clinically significant. We are seeking an independent opinion on this point and may be in touch at a later date with further questions.

5 Response:
See response 4.4

References

In addition we have provided a detailed biography of the current progress of the project to date. This bibliography provides further details of areas of uncertainty.

1. Jefferson T, Jones M, Doshi P, Del Mar C. Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis. BMJ 2009;339:b5106

2. Doshi P. Neuraminidase inhibitors: the story behind the Cochrane review. BMJ 2009;339:b5164

3. Cohen D. Complications: tracking down the data on oseltamivir. BMJ 2009;339:b5387.

4. Godlee F, Clarke M. Why don't we have all the evidence on oseltamivir? BMJ 2009;339:b5351
5. Editor's Choice: We want raw data, now. BMJ 2009; 339 doi: 10.1136/bmj.b5405

6. Smith J, on behalf of Roche. Point-by-point response from Roche to BMJ questions. BMJ 2009;339:b5374.

7. Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database of Systematic Reviews 2012, Issue 1. Art. No.: CD008965. DOI: 10.1002/14651858.CD008965.pub3.

8. Doshi P, Jones MA, Jefferson T. Rethinking credible evidence synthesis. BMJ 2012;344:d7898 doi: 10.1136/bmj.d7898.

9. Drug Data Shouldn't Be Secret by Peter Doshi and Tom Jefferson
   The New York Times, April 10, 2012. URL: http://www.nytimes.com/2012/04/11/opinion/drug-data-shouldnt-be-secret.html Shortened URL: http://nyti.ms/Ivgh9c

10. Doshi P, Jefferson T, Del Mar C (2012) The Imperative to Share Clinical Study Reports: Recommendations from the Tamiflu Experience. PLoS Med 9(4): e1001201. doi:10.1371/journal.pmed.1001201
    Short URL: http://bit.ly/HIbwqO. PDF for printing: http://bit.ly/HFBYTV

11. Eichler H-G, Abadie E, Breckenridge A, Leufkens H, Rasi G (2012) Open Clinical Trial Data for All? A View from Regulators. PLoS Med 9(4): e1001202. doi:10.1371/journal.pmed.1001202

12. Jones M, Hama R, Jefferson T, Doshi P. Neuropsychiatric Adverse Events and Oseltamivir for Prophylaxis Drug Saf 2012; 35 (12): 1187-1190.

13. Tamiflu open data campaign
    http://www.bmj.com/tamiflu

14. Godlee F. Clinical trial data for all drugs in current use. BMJ 2012;345:e7304 doi: 10.1136/bmj.e7304 (Published 29 October 2012)

15. http://www.bmj.com/content/345/bmj.e7304

16. Payne D. Tamiflu: the battle for secret drug data. BMJ 2012;345:e7303 doi: 10.1136/bmj.e7303 (Published 29 October 2012) http://www.bmj.com/content/345/bmj.e7303

17. Zosia Kmietowicz. Academics plea for politicians to tackle problem of missing data. BMJ 2012;345:e7306 doi: 10.1136/bmj.e7306 (Published 29 October 2012) http://www.bmj.com/content/345/bmj.e7306

18. Godlee F. Open letter to Roche about oseltamivir trial dataBMJ 2012;345:e7305 doi: 10.1136/bmj.e7305 (Published 29 October 2012). http://www.bmj.com/content/345/bmj.e7305

19. http://www.nytimes.com/2012/11/01/business/british-medical-journal-to-require-detailed-clinical-trial-data.html?_r=0

This document has been authored and approved by the Cochrane NI team and signed on behalf by

Professor Carl Heneghan BM, BCH, BA, MA, MRCGP, DPhil (Oxon)

Director of the Centre for Evidence-Based Medicine, Clinical Reader
& Fellow of Kellogg College, University of Oxford
Email: carl.heneghan@phc.ox.ac.uk

Appendix 6A Anonymised peer review comments

Thursday, 06 February 2014

Contact editor review comments

STRUCTURE OF THE REPORT

This is a huge piece of work – many congratulations! There are actually eight (2x2x2) comparisons in this one review, two drugs, treatment and prophylaxis (and then adults and children for each of these). This makes the structure of your report challenging. At present it reads like a woven rug – threads of the results from each of the above are interwoven and hard to untangle for the reader. It would be hugely helpful to rationalise the structure and put in headings throughout. At present it does not work to use the outcomes as the headings for the narrative, and then try to describe these eight comparisons under each outcome! This is particularly because the primary outcomes are not the same for the treatment and prophylaxis comparisons. You will need to decide how to guide the reader through this maze but I have some practical suggestions below:

SUMMARY OF FINDINGS TABLES

It is great to see the four existing SoF tables in the review. These are really helpful with such a huge body of evidence. However they could be improved (in my view) in the following ways, as they do not currently match the primary outcomes in the body of the review:

1. I suggest **separate tables for adults and children** (so there would be 8 SoF tables). This will allow you to put a line in for *all the primary outcomes* in each comparison and show more data for the outcomes in children.
2. So for **treatment tables** there should be a line for time to first alleviation of symptoms (as there is at present, but add the mean duration on control in the empty first box). Could you make both drugs into hours or days to aid the reader here?
3. The treatment studies should then have a line for hospitalisation (even if there is no data viz zanamivir, put in a comment to say this is missing).
4. Then two lines for complications (as at present for pneumonia) but maybe bronchitis as well.
5. That will leave three lines for adverse events: maybe vomiting, diarrhoea and cardiac for the treatment comparisons.
6. The **prophylaxis tables** should have lines for symptomatic and asymptomatic influenza (as at present), and then hospitalisations (even if empty) with the remains lines for adverse events. However please check the **units of analysis for the psychiatric complications** before they go into the SoF (see results below).
7. **Please add the GRADE of evidence in a final column to the SoF tables**? You have assessed risk of bias so you could use it here!
8. There are some **errors in the current SoF table 4**: the number of participants in the first two existing rows is wrong. I make it 5275 (4) not 5976 (7), and the RR in the first row should be 0.39 not 0.45 (although the RD looks correct). Also asymptomatic lab-confirmed influenza seems to have lost its subgroups in the Forest plot 4.4 (the RR in the SoF looks to be correct though).
9. I do not think that you need the median risk line as well as the mean risk in these tables.

ABSTRACT

Please add headings to **the results section** to help the reader pick out what outcomes are being considered, and whether it is treatment or prophylaxis (see bold text in screen shot below).

Main results

For treatment of adults, oseltamivir reduced the **time to first alleviation of symptoms** by 16.8 hours (95% confidence interval (CI) 8.4 to 25.1 hours, P < 0.0001). This represents a reduction in the time to first alleviation of symptoms from 7 to 6.3 days. In children there was no significant effect on time to first alleviation of symptoms, including those with asthma. Zanamivir reduced time to first alleviation of symptoms in adults by 0.60 days (95% CI 0.39 to 0.81 days, P < 0.00001) equating to a reduction in the mean duration of symptoms from 6.6 to 6.0 days. The effect in children was not significant.

Treatment of adults with oseltamivir had no effect on **hospitalisations** risk difference (RD) = 0.15% (95% CI -0.91 to 0.78, P = 0.84), with insufficient evidence in children and in prophylaxis (zanamivir hospitalisation data: unreported).

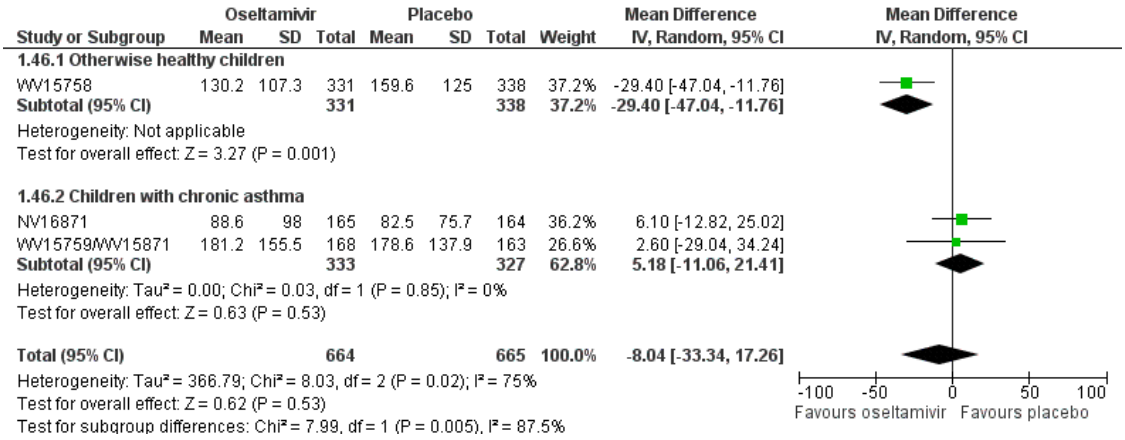
In adult treatment trials, there was no evidence that oseltamivir reduced those **complications** classified as serious or which led to study withdrawal RD = 0.07% (95% CI -0.78 to 0.44). Oseltamivir reduced self-reported investigator-mediated unverified pneumonia, RD = 1.00% (95% CI 0.22 to 1.49); number needed to benefit (NNTB) = 100 (95% CI 67 to 451) in the treated population. The effect was not significant in the five trials that used a more detailed diagnostic form for pneumonia. There were no definitions of pneumonia (or other complications) in any trial. No oseltamivir treatment studies reported effects on radiologically-confirmed pneumonia. There was no evidence of effect on unverified pneumonia in children. The was no evidence of an effect of zanamivir in either self reported or radiologically-confirmed pneumonia In prophylaxis trials. In prophylaxis zanamivir reduced the risk of self-reported pneumonia in adults, RD = 0.32% (95% CI 0.09 to 0.41); NNTB = 311 (95% CI 244 to 1086) but not oseltamivir. Zanamivir reduced the risk of bronchitis in adult treatment trials, RD = 1.80% (95% CI 0.65 to 2.80); NNTB = 56 (36 to 155) but not oseltamivir. There was no evidence that either NI reduced the risk of otitis media and sinusitis in both adults and children.

Oseltamivir in treatment of adults was associated with **increased risk of nausea**, RD = 3.66% (95% CI 0.90 to 7.39); number needed to treat to harm (NNTH) = 28 (95% CI 14 to 112) and vomiting, RD = 4.56% (95% CI 2.39 to 7.58); NNTH = 22 (14 to 42). Oseltamivir was associated with decreased risk of diarrhoea RD = 2.33%; (95% CI 0.14 to 3.81); NNTB = 43 (95% CI 27 to 709) and cardiac events, RD = 0.68% (95% CI 0.04 to 1.0), NNTB = 148 (101 to 2509), compared to placebo during on-treatment period. In treatment of children, oseltamivir induced vomiting, RD = 5.34% (95% CI 1.75 to 10.29); NNTH = 19 (95% CI 10 to 57).

In **prophylaxis trials** oseltamivir and zanamivir reduced the risk of symptomatic influenza in individuals and households. There was no evidence of effect on asymptomatic influenza or on non-influenza influenza-like illness. In oseltamivir prophylaxis studies, psychiatric adverse events were increased in the combined on- and off-treatment periods, RD = 1.25% (95% CI 0.19 to 3.05); NNTH = 81 (95% CI 33 to 540) in the study treatment population. Data suggested a dose-response effect on psychiatric events in the two oseltamivir "pivotal" treatment trials WV15670 and WV15671 at 150 mg (standard dose) and 300 mg daily (high dose) (P = 0.038). In prophylaxis studies, oseltamivir increased risk of headaches on-treatment, RD 3.15% (95% CI 0.88 to 5.78); NNTH = 32 (95% CI 18 to 115), renal events on treatment, RD 0.67% (95% CI -2.93 to 0.01); NNTH = 150 (NNTH 35 to NNTB > 1000) and nausea whilst on-treatment, RD = 4.15%, (95% CI 0.86 to 9.51); NNTH = 25 (95% CI 11 to 116).

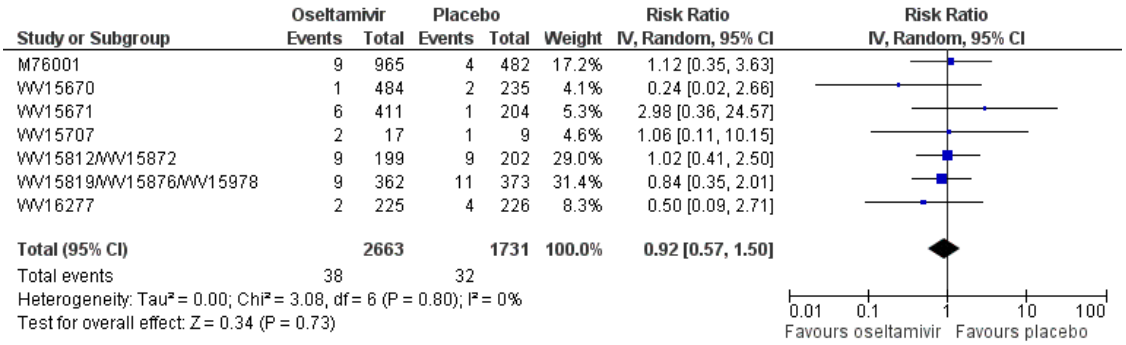
The highlighting in yellow is to point out that **the way that results are described is not uniform in the abstract**. Ideally pick hours or days for both drugs (I guess this is a hang-over from the combination of two previous reviews)?

I am very uncomfortable about the diverse wording to describe the results. For example the MD of the time to first alleviation of symptoms is not stated in children, and the heterogeneity is not described. It was not statistically significant, I agree, but this is somewhat selective in reporting this outcome. I have included the random effects model in the Forest plot below (although in the current version of the review this seems to be a fixed model for some reason could be my fault but please check)? There is very large heterogeneity and a significant difference between the results in asthmatic children and healthy children. Overall the pooled 95% CI overlaps with the adults, so I believe the current reporting is misleading in that we cannot rule out the same benefit in healthy children and adults……



"**No effect on hospitalisations**" is also troubling. Have you pre-defined the boundaries of the 95% CI that you need to rule out an important effect? I guess not as this did not appear in the methods. You cannot say "no effect" unless you set boundaries and find a 95% CI that is within those boundaries, in order to rule out important benefit or harm. The Forest plot below shows **a fairly wide 95% CI for the RR**. I suggest that we cannot rule out an effect on

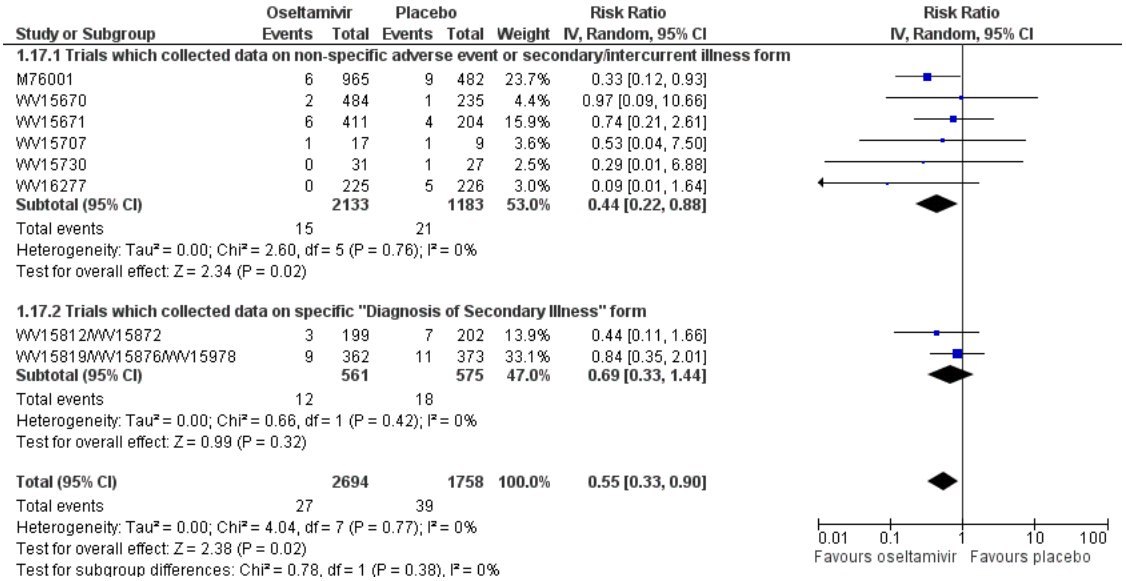
hospitalisations from the Forest plot. **Also why does the P value in the text not match that in the Forest plot? Also see section on P values in relation to results reporting later.**



Later on you use “no evidence of an effect” which is technically more correct, but very hard to interpret without a point estimate and confidence intervals. Similarly “not oseltamivir” is uninformative. **Although the RD and NNT are very helpful, they are hard to ground without any mention of control event rate or RR in the text of the abstract or the results in the full review.**

This is much clearer in the SoF tables, as full data is included. The 1 percentage point reduction in risk of pneumonia needs to be set in the context of a 2% risk overall on placebo, to show that the risk is halved.

This comment highlights a difficulty that repeats itself throughout the reporting of the results (as they are reported as RD and NNT with 95% CI, but not control event rates, and missing number of participants and studies) in the text of the review too. Could this be addressed please?



MISSING ESTIMATE OF THE SIZE OF THE REDUCTION IN RISK OF SYMPTOMATIC INFLUENZA

In the prophylaxis trials no idea is given of the size of reduction in the risk of **symptomatic influenza**. I am sure this is nothing to do with the fact that these are some of the largest relative and absolute treatment effects in the review! Please **report RD and NNT for these outcomes** in the same way as you do for the adverse events. How else can the reader trade them off?

FOREST PLOTS

Could these be re-ordered to put the primary outcomes at the top?

Comparison 4.4 has lost its subgroups.

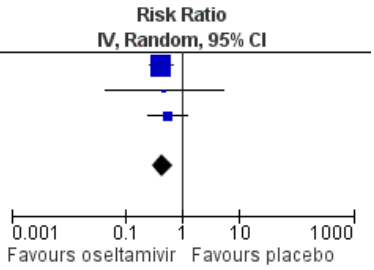
If you decide to use the P values and I^2 results from the Forest plots, please check that they are consistent with the text throughout (see results section below).

REPORTING OF RESULTS IN THE MAIN TEXT

As above in the result reporting in the main text the result of symptomatic influenza in prophylaxis trials is described as "marginally lower" with no point estimate or 95% CI of any kind…. This is not adequate reporting of the results from this outcome (Forest plot and three Cates plots for Oseltamivir prophylaxis then Zanamivir then nausea on Oseltamivir for comparison) and needs to be amended please.

Study or Subgroup	Oseltamivir Events	Total	Placebo Events	Total	Weight	Risk Ratio IV, Random, 95% CI
WV15673/WV15697	30	1040	36	519	72.0%	0.42 [0.26, 0.67]
WV15708	1	190	2	182	2.8%	0.48 [0.04, 5.24]
WV15825	9	276	16	272	25.2%	0.55 [0.25, 1.23]
Total (95% CI)		**1506**		**973**	**100.0%**	**0.45 [0.30, 0.67]**
Total events	40		54			

Heterogeneity: Tau² = 0.00; Chi² = 0.37, df = 2 (P = 0.83); I² = 0%
Test for overall effect: Z = 3.91 (P < 0.0001)



Below is a screen shot of this pooled Forest plot data entered into Visual Rx on my website, which produces the event rate for each group and NNTB with 95% CI. The next page shows the resulting Cates plot for this outcome. In my view the reader needs to see this data and make up their own mind whether this is marginal or not.

Input: HELP

Intervention Title: Oseltamivir for household contacts
Outcome: Lab confirmed symptomatic influenza
Duration: some weeks
Control Event %: 5.55
Relative Risk: 0.45
Lower CI: 0.30
Upper CI: 0.67
Adverse events: ☑
Number of faces: 1000
Method: Relative Ris

Calculate | Reset | Draw Cates plot

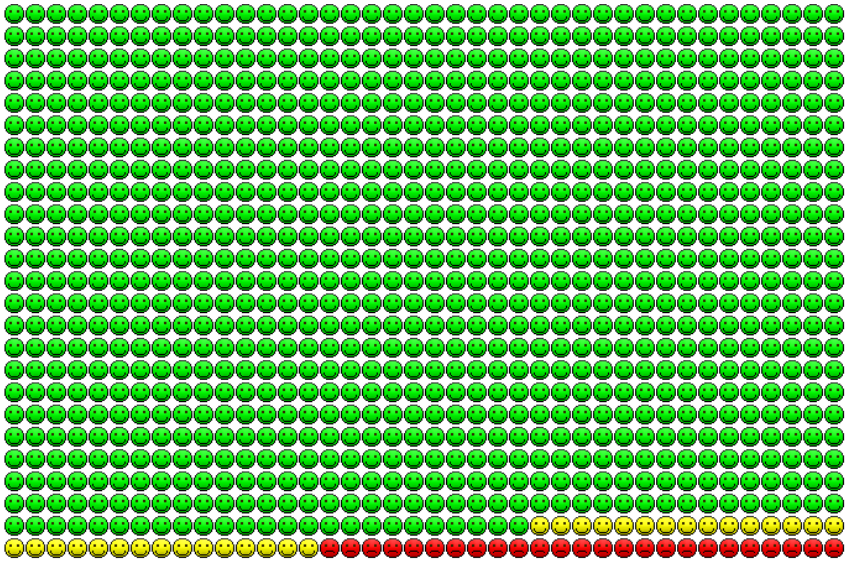
Output:

Oseltamivir for household contacts

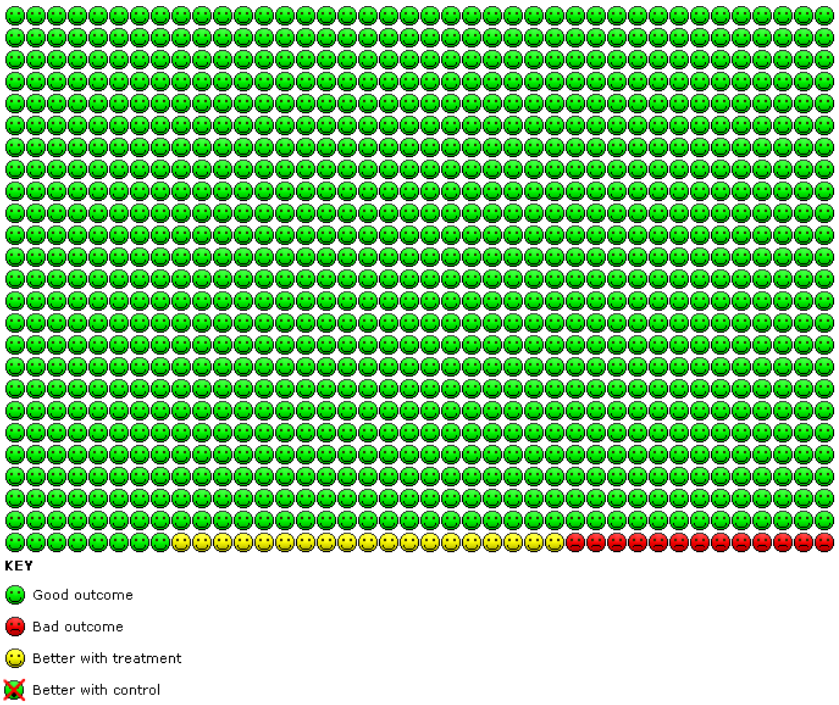
Outcome: Lab confirmed symptomatic influenza
Duration: some weeks

Control group risk	Treatment group risk (95% CI)	NNTB (95% CI)
5.55%	2.50% (1.66% to 3.72%)	33 (NNTB 26 to NNTB 55)

In the control group 56 people out of 1000 had Lab confirmed symptomatic influenza over some weeks, compared to 25 (95% CI 17 to 37) out of 1000 for the active treatment group.



Cates plot of symptomatic influenza in household contacts with Oseltamivir



Cates plot of symptomatic influenza with Zanamivir prophylaxis.

Input: HELP

Intervention Title:	Oseltamivir for treatment
Outcome:	Nausea
Duration:	some weeks
Control Event %:	6.40
Relative Risk:	1.57
Lower CI:	1.14
Upper CI:	2.15
Adverse events:	☑
Number of faces:	1000
Method:	Relative Ris

Calculate | Reset | Draw Cates plot

Output:

Oseltamivir for treatment

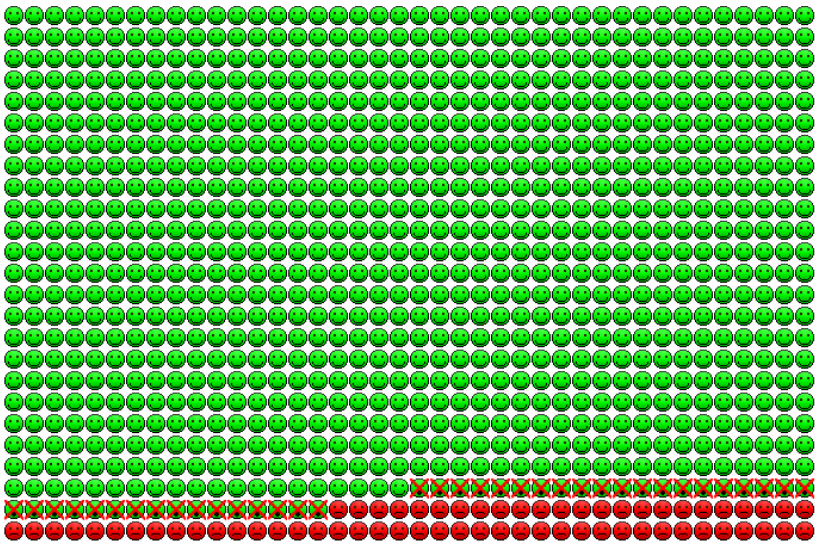
Outcome: Nausea

Duration: some weeks

Control group risk	Treatment group risk (95% CI)	NNTH (95% CI)
6.40%	10.05% (7.30% to 13.76%)	28 (NNTH 112 to NNTH 14)

In the control group 64 people out of 1000 had Nausea over some weeks, compared to 100 (95% CI 73 to 138) out of 1000 for the active treatment group.

Data entry screen for nausea on Oseltamivir shown above from pooled RR and Cates plot below



KEY

Good outcome

Bad outcome

Better with treatment

Better with control

Cates plot of nausea on Oseltamivir treatment (for comparison with the prophylaxis Cates plots above)

Also I could not tell how you decided to describe results as "no evidence" or "insufficient evidence", as the 95% CI for the latter is narrower than the former in the section on serious complications in the last paragraph below.

Analysis of hospitalisations

In oseltamivir treatment of adults there was no evidence of a difference in hospitalisation rate between treatment groups, RD = 0.15% (-0.91 to 0.78) (Figure 8). There was insufficient evidence to show an effect on hospitalisations in children (Figure 9) and in prophylaxis (Figure 10). Data on hospitalisations for the zanamivir studies were not reported.

Analysis of influenza complications

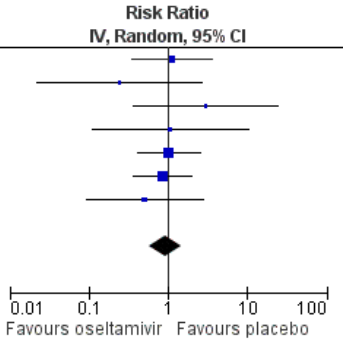
In adult treatment trials, oseltamivir reduced self-reported investigator-mediated unverified pneumonia, RD = 1.00% (95% CI 0.22 to 1.49, P = 0.05, I^2 statistic = 0%); NNTB = 100 (95% CI 67 to 451) in the treated population. The effect was not significant in the five trials that used a more detailed diagnostic data collection form and no studies reported on radiological confirmation of pneumonia. (Figure 11) There was no evidence of an effect on pneumonia in children (Figure 12). In two zanamivir adult trials (NAI30012; NAI30015) pneumonia reporting was based on a stricter definition of X-ray confirmation and there was no evidence of a treatment effect (Figure 13). In nine zanamivir trials (NAI30008; NAI30010; NAI30011; NAIA/B2008; NAIA2005; NAIA3002; NAIB2007; NAIB3001; NAIB3002) pneumonia was a self-reported investigator-mediated unverified outcome (Figure 14; Figure 15). Overall there was no evidence of an effect of zanamivir on mixed verified and unverified pneumonia in adult treatment (Analysis 3.2).

In oseltamivir trials there was no evidence that treatment affected complications classified as serious or that led to withdrawal from the trial, RD = 0.07% (95% CI -0.78 to 0.44) (Analysis 1.20; Analysis 1.55 in adults and children). This outcome could not be assessed in oseltamivir prophylaxis due to insufficient number of events. There was insufficient evidence to show zanamivir in adult treatment and prophylaxis reduced the risk of any complication classified as serious or which led to study withdrawal, RD = -0.04% (95% CI -0.64 to 0.24) (Analysis 3.7; Analysis 4.8) and this outcome could not be assessed in children due to insufficient number of events.

I can understand how you might apply this to the hospitalisations, as there were only 18 events in children for this outcome (but did you use an agreed threshold for this)? At present any distinction between these two descriptions is muddled in the reporting of results. **Please revisit the wording in the results.**

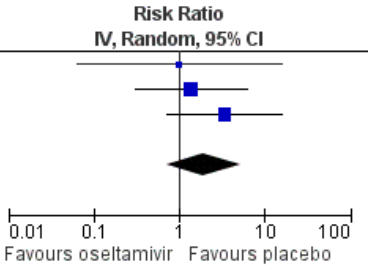
Study or Subgroup	Oseltamivir Events	Total	Placebo Events	Total	Weight	Risk Ratio IV, Random, 95% CI
M76001	9	965	4	482	17.2%	1.12 [0.35, 3.63]
WV15670	1	484	2	235	4.1%	0.24 [0.02, 2.66]
WV15671	6	411	1	204	5.3%	2.98 [0.36, 24.57]
WV15707	2	17	1	9	4.6%	1.06 [0.11, 10.15]
WV15812/WV15872	9	199	9	202	29.0%	1.02 [0.41, 2.50]
WV15819/WV15876/WV15978	9	362	11	373	31.4%	0.84 [0.35, 2.01]
WV16277	2	225	4	226	8.3%	0.50 [0.09, 2.71]
Total (95% CI)		**2663**		**1731**	**100.0%**	**0.92 [0.57, 1.50]**
Total events	38		32			

Heterogeneity: Tau² = 0.00; Chi² = 3.08, df = 6 (P = 0.80); I² = 0%
Test for overall effect: Z = 0.34 (P = 0.73)



Study or Subgroup	Oseltamivir Events	Total	Placebo Events	Total	Weight	Risk Ratio IV, Random, 95% CI
NV16871	1	165	1	164	13.2%	0.99 [0.06, 15.76]
WV15758	4	342	3	353	45.3%	1.38 [0.31, 6.10]
WV15759/WV15871	7	170	2	165	41.5%	3.40 [0.72, 16.11]
Total (95% CI)		**677**		**682**	**100.0%**	**1.92 [0.70, 5.23]**
Total events	12		6			

Heterogeneity: Tau² = 0.00; Chi² = 0.93, df = 2 (P = 0.63); I² = 0%
Test for overall effect: Z = 1.27 (P = 0.20)



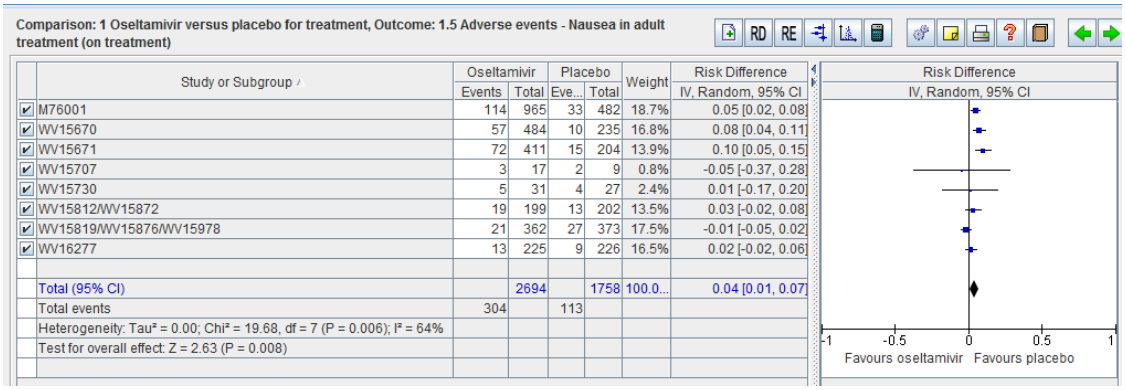
P VALUES AND I^2 REPORTING FOR RD AND 95% CI IN THE TEXT OF THE RESULTS

P values and I^2 stats are reported in some places in the results with the RD and its 95% CI. At least in one instance that I found this *appears* to be derived from an analysis of **pooled risk difference! This is not the methodology that**

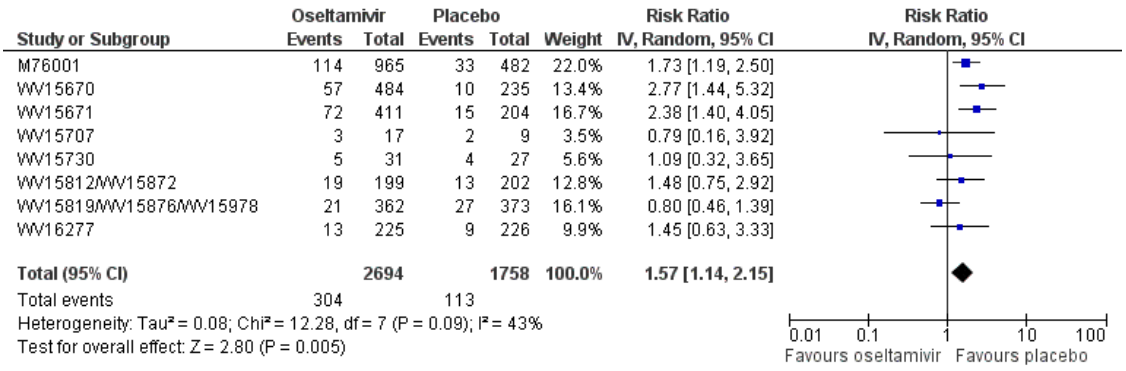
you specify in the methods or show in the Forest plots.

Analysis of harms

Oseltamivir in treatment of adults is associated with increased risk of nausea, RD = 3.66% (95% CI 0.90 to 7.39, P = 0.008, I^2 statistic = 64%); NNTH = 28 (95% CI 14 to 112) (Figure 27) and vomiting, RD = 4.56% (95% CI 2.39 to 7.58, P =0.001, I^2 statistic = 75%); NNTH = 22 (95% CI 14 to 42) (Figure 28) and decreased risk of diarrhoea, RD = 2.33% (95% CI 0.14 to 3.81, P = 0.04, I^2 statistic = 42%); NNTB = 43 (95% CI 27 to 709) (Figure 29), compared to placebo during on-treatment periods. Both of these outcomes were associated with significant heterogeneity where treatment effects appeared larger in



The picture above is my **RD analysis** and the one below is the **RR analysis** in Figure 27.



At the least the derivation of P values and I^2 needs to be made clear to the reader (from the original RR analysis or somewhere else)? They do not really apply to the RD and 95% CI that is reported in the text. My suggestion would be to take these P values and I^2 figures out as there is huge potential for confusion at present. If you leave them in, then you will need to explain how they are derived in the methods section.

PSYCHIATRIC ADVERSE EVENTS

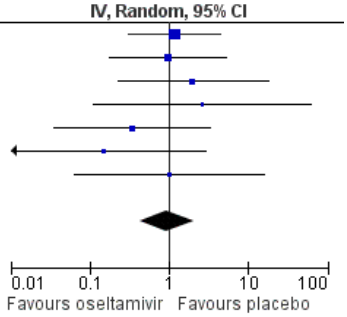
I could not make any sense of the comment (yellow highlight in final para below) in relation to Peto Odds ratio and the P value of 0.06 presented as this bears no relation to the P value in Figure 31:

Analysis of harms

Oseltamivir in treatment of adults is associated with increased risk of nausea, RD = 3.66% (95% CI 0.90 to 7.39, P = 0.008, I^2 statistic = 64%); NNTH = 28 (95% CI 14 to 112) (Figure 27) and vomiting, RD = 4.56% (95% CI 2.39 to 7.58, P =0.001, I^2 statistic = 75%); NNTH = 22 (95% CI 14 to 42) (Figure 28) and decreased risk of diarrhoea, RD = 2.33% (95% CI 0.14 to 3.81, P = 0.04, I^2 statistic = 42%); NNTB = 43 (95% CI 27 to 709) (Figure 29), compared to placebo during on-treatment periods. Both of these outcomes were associated with significant heterogeneity where treatment effects appeared larger in otherwise healthy adults compared to the elderly and the chronically ill. However one trial of otherwise healthy adults (WV16277) also showed smaller effects. The cardiac effects of oseltamivir are unclear. Exposure to oseltamivir may reduce cardiac general events compared to placebo, RD = 0.68% (95% CI 0.04 to 1.00, P = 0.11, I^2 statistic = 0%); NNTB = 148 (101 to 2509) (excluding WV16277 in which ECG was included in the safety parameters) (Figure 30) but may increase QTc prolongation (including borderline) in trial WV16277 (RD 4.0% (0.71 to 7.30), NNTB = 25, 14 to 140) compared to placebo during on-treatment periods.

In treatment trials there is insufficient evidence to show an association between oseltamivir and on-treatment psychiatric adverse events overall (Figure 31). However, there was a dose-response effect in the two "pivotal" treatment trials. In identically designed trials WV15670 and WV15671 there were two active treatment groups: 75 mg (standard dose) and 150 mg (high dose) oseltamivir per day. In dose response analysis there was evidence of increased risk of psychiatric body system adverse events over the entire follow up period (P = 0.038 based on likelihood ratio test). In trial WV15670 the event rates were: 1/204; 1/206; 4/205 in placebo, 75 mg and 150 mg arms respectively, whereas trial WV15671 had rates of 2/235; 0/242; 5/242 respectively. Figure 32 shows that in prophylaxis trials of oseltamivir there was a significant increase in psychiatric adverse events over the on and off-treatment periods, RD = 1.25% (95% CI 0.19 to 3.05, P = 0.04, I^2 statistic = 4%); NNTH = 81 (95% CI 33 to 540). Initial analysis of patients with psychiatric adverse events in the on-treatment period showed a borderline statistically significant result (P = 0.06) hence we conducted sensitivity analysis using Peto's method (P = 0.05) as well as the analysis reported in Figure 31.

	Oseltamivir		Placebo			Risk Ratio	Risk Ratio
Study or Subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**IV, Random, 95% CI**	**IV, Random, 95% CI**
M76001	7	965	3	482	33.3%	1.17 [0.30, 4.49]	
WV15670	4	484	2	235	21.2%	0.97 [0.18, 5.26]	
WV15671	4	411	1	204	12.7%	1.99 [0.22, 17.65]	
WV15730	1	31	0	27	6.1%	2.63 [0.11, 61.88]	
WV15812/WV15872	1	199	3	202	11.9%	0.34 [0.04, 3.23]	
WV15819/WV15876/WV15978	0	362	3	373	6.9%	0.15 [0.01, 2.84]	
WV16277	1	225	1	226	7.9%	1.00 [0.06, 15.96]	
Total (95% CI)		**2677**		**1749**	**100.0%**	**0.93 [0.43, 2.03]**	
Total events	18		13				
Heterogeneity: Tau² = 0.00; Chi² = 3.25, df = 6 (P = 0.78); I² = 0%							
Test for overall effect: Z = 0.18 (P = 0.86)							



I have a larger concern about the way the psychiatric adverse events were gathered, and whether you know for sure that each event was in a different person. My experience is that the same person can have a single adverse event recorded under multiple labels. **If you know that the numbers in the Forest plots are individual people (from the trial reports) please say so in the results and in the methods under unit of analysis issues. If we do not know this for sure then the Forest plots are not valid, as there may be multiple counting of the same event in one person under different headings and you will need to reconsider this outcome.**

RISE IN ANTIBODIES

This is currently described under analysis of harms, could you create a new section on antibodies to accommodate these results?

DISCUSSION

Usually starts with a summary of the main results?

Please add sub-headings to help the reader find their way around.

Please search for “no effect” and see whether you really mean this!

What is SBA or JSBA?

“strangely believed” might be better reworded in para following Fritz 1999 reference.

I wonder if animal studies should be in an appendix or left out (cf Peer reviewer).

Can you define what you mean by **“no credible effect”** of oseltamivir against pneumonia, or re-word it? Similarly you use the term credible evidence in the conclusions.

I would love to see a summary of the differences in the findings on **treatment effects** in this review compared to the previous version, and compare to Hernan and Litsch, not just the differences in methodology.

AUTHORS’ CONCLUSIONS

Meaning of credible evidence is not clear. I do not think that your statement of “**minimal effect**” of NIs at prophylactic treatment is justified by the full data in the review. These outcomes had the smallest NNTs in the review, which are not reported in the text! **Please amend this.**

⊟ Authors' conclusions

⊟ Implications for practice

On the basis of the findings of this review, clinicians and health care policy makers should urgently revise current recommendations for use of the NIs for individuals with influenza. Our findings confirm that both oseltamivir and zanamivir reduce the time to symptomatic improvement in adults (but not asthmatic children) with influenza-like illness. The size of this effect is small, approximately half a day. It is unclear whether this is superior to treatment with commonly used antipyretic medications. However, we did not find any credible evidence that either oseltamivir or zanamivir reduce risk of complications of influenza, particularly pneumonia, nor reduce risk of hospitalisation or death. Moreover, even in individuals at higher risk of complications, such as children with asthma, or the elderly, we found no evidence of beneficial effect for reducing risks of complications.

Based on these findings there appears to be no evidence for patients, clinicians or policy makers to use these drugs to prevent serious outcomes, both in annual influenza and pandemic influenza outbreaks. Practice recommendations and drug labelling needs to be changed to reflect these findings.

When used as prophylactic agents to prevent the occurrence of influenza in individuals or families, our findings again suggest a minimal effect and based on this there is little support for their use as prophylactic agents for example during influenza epidemics. Given that oseltamivir is now recommended as an essential medicine for treatment of seriously ill patients or those in higher-risk groups (WHO 2013a; WHO 2013b) with influenza (H1N1 2009), this is of some concern.

Perhaps rephrase to suggest that the WHO “revisit the evidence following your work, as the guidelines may need revision”? It is up to them whether they revise their guidance. This is beyond your brief! Please revisit

IMPLICATIONS FOR RESEARCH

I think this is too strong! You raise uncertainties in the body of the text (insufficient evidence), particularly in relation to sparse data and uncertainty of treatment effects in children, and then appear to ignore these uncertainties in your recommendation to discontinue current trials! Please revisit

MINOR POINTS

I found duplicate sentences in relation to Cochrane Neuraminidase Inhibitors review team in the discussion, and Toby Lasserson in the acknowledgements!

Consumer peer review comments

Cochrane Intervention Review, consumer peer-referee form

Title:	**A151** - Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children
Name of referee:	[redacted]
Date to be returned to editorial base:	31 January 2014

1. TITLE

Please note that the title has already been peer-reviewed and agreed to in the Cochrane Protocol (available at www.thecochranelibrary.com*). We therefore we do not expect you to comment on the title. If you can suggest an improvement, however, please do so here and the editors will consider it.
Okay

2. ABSTRACT

Is there anything mentioned in the Cochrane Review that may be important, but is missing from the Abstract? Do you think that the Abstract overstates or understates what was found in the Cochrane Review?

Does not mention the FDA data so suggest add something like:

Data collection and analysis
We obtained 107 clinical study reports from the European Medicines Agency, GlaxoSmithKline and Roche. We accessed FDA documents from the internet.

3. PLAIN LANGUAGE SUMMARY (PLS)

Suggested changes:

A review of ~~unpublished~~ the regulatory information on~~from~~ trials of the neuraminidase inhibitors ~~(Tamiflu -~~ oseltamivir (Tamiflu) and zanamivir (Relenza~~ zanamivir~~) for influenza in adults and children

Oseltamivir and zanamivir are classified by the World Health Organization as essential medicines, and have been stockpiled in many countries to treat and prevent seasonal and pandemic influenza, before a 'flu vaccine' matched to the contributing virus becomes available.

How this review has been approached
We ~~decided to~~ have updated and combined/~~amalgamate~~d our reviews on the antiviral drugs zanamivir and oseltamivir for influenza in adults and children on the basis of the manufacturers' reports to regulators (~~called~~ clinical study reports) and the regulators' comments (which together? we called regulatory information). Clinical study reports are unpublished extensive documents with great~~exhaustive~~ detail~~s~~ on~~f~~ the trials that formed the basis for market approval. They include the protocols, methods and results. Clinical study reports have hitherto been confidential, seen only by manufacturers and ~~in part by~~ regulators.

Why we have taken this approach

In previous versions of this review we have identifiedview of the unresolved discrepancies in the data presented in published trial reports and of the substantial publication bias., wWe therefore elected not to use data from journal articles but addressed the. Aavailability of documents generated by regulatory bodies during licensing processes. We have accessed such data fromin the UK, USA, European Medicines Agency (EMA) and Japan; and clinical study reports from the manufacturers (after a protracted media campaign)., This has enabled us to verify information from the randomised placebo-controlled trials on adults and children with confirmed or suspected exposure to naturally occurring influenza.
Based on our assessments of the regulatory documents (in excess of 160,000 pages), we have caome to the conclusion that there were substantial problems with the design, conduct, reporting and availability of information from many of the trials.

What we have found

We found that both drugs shorten the durations of symptoms of influenza-like illness (unconfirmed influenza, or (‘"the flu’") by less than a day. but oOseltamivir didoes not affect the number of hospitalisations based on the data from all the people enrolled in treatment trials of oseltamivir; z. Zanamivir trials did not record this outcome. The effects on pneumonia and other complications of influenza such as bronchitis, middle ear infection (otitis media) and sinusitis were unreliably recorded, as shown by the case report form in the trial documents that had limitations in diagnostic criteria for pneumonia and missing follow up of diary cards from participants. In children with asthma there was no clear effect on time to first alleviation of symptoms.
Prophylaxis trials showed that oseltamivir and zanamivir reduced the risk of symptomatic influenza in individuals and households. There was no evidence of effect on asymptomatic influenza or on non-influenza influenza-like illness. Oseltamivir use was associated with nausea, vomiting, headaches, renal and psychiatric events, these last three when used to prevent influenza (prophylaxis); it may aid heartbeat regularity. In adult treatment trials of zanamivir there was no increased risk of reported adverse events. The evidence was insufficient to show harms associated with treatment of children with zanamivir.

Agreement with other findings

In this lack of effect on complications, our independent analysis concurs with the conservative conclusions on the effects onf both drugs by the US Food and Drug Administration (FDA). The FDA only allowed claims of effectiveness of both drugs for the prevention and treatment of symptoms of influenza and not on other effects (includingsuch as the interruption of person-to-person spread of the influenza virus or prevention of pneumonia). The FDA described the overall performance of both drugs as ‘"modest’". There is evidence to suggest that oseltamivir use is associated with nausea, vomiting, headaches, renal and psychiatric events (these last three when used to prevent influenza/ Prophylaxis is the mode of use of NIs when there is expectation of possible near future exposure to influenza, that is in a prophylactic role) and perhaps raised? blood sugar levels but may aid heartbeat regularity.

Mechanism of action for beneficial effects

These findings all suggest that the low immune response with low levels of pro-inflammatory cytokines induced by the action
of oseltamivir carboxylate may reduce symptoms of influenza unrelated to inhibition of influenza virus replication. The potential hypothermic or antipyretic effect of oseltamivir as a central nervous system depressant (but not zanamivir) may also contribute to the apparent reduction of host symptoms. Statements made on the capacity of oseltamivir to interrupt viral transmission and reduce complications are not supported by any data we have been able to access.
The mechanism of action proposed by the producers (influenza virus-specific), does not fit the evidence. The evidence suggests a direct effect of oseltamivir on many systems and on antibody production, perhaps mediated by an anti-inflammatory action.

CONTENT OF THE PLS
Does the title of the PLS reflect the title of the Cochrane Review, and is it easy to understand? If not, can you identify which words or phrases are difficult to understand, or could you suggest any improvements to the wording? **Made more explicit**

Is the health problem or issue being addressed stated clearly? **Yes**
Are the interventions and comparisons/controls examined in the Cochrane Review stated clearly and succinctly in this section? **Yes**
Does the PLS report the main findings from the Cochrane Review clearly and accurately? Does it report on adverse effects or harms? **Yes**
Does the PLS describe the overall quality of the evidence, and comment on any issues that could affect the findings of the review?
Do you think the findings in the PLS are consistent with the Abstract and the rest of the Cochrane Review? Is there anything mentioned in the Review that may be important, but is missing from the PLS? Do you think that the PLS overstates or understates what was found in the Review? **Made some suggestions above**
Do you think the PLS would help patients, carers and the public in making a healthcare decision? If not, is there anything missing from the PLS that you think should be included? Do you have any other suggestions for improvement? **I have split the summary up with headings that explain and prepare the reader for the content**
WRITING STYLE OF THE PLS
Is the PLS written in plain language and easy to understand? Are sentences too long or wordy? Are there any parts that you think should be rewritten? **Yes**
Are abbreviations, research terms and technical terms avoided or explained? **YES, largely so**

4. BACKGROUND/OBJECTIVES/CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW/SEARCH METHODS FOR IDENTIFICATION OF STUDIES/DATA COLLECTION AND ANALYSIS

Please note that these sections were published in the Cochrane Protocol (available at www.thecochranelibrary.com*) and have therefore been peer-reviewed. If you would like to provide any comments, please do so here and the editors will consider them. Major suggestions for change are more likely to be considered for future updates of the Cochrane Review than this version. **Excellent background to the review**

5. RESULTS

Can you understand the format of the results? Is it clear whether the intervention was effective or not?

Yes clearly reported and well split up under headings
Do the results include information about the overall quality of the evidence, and risk of bias? **Yes**
Does the section 'Included studies' and the 'Characteristics of included studies' table include details about the funding sources for the studies?

6. DISCUSSION

Can you identify any words or phrases that are difficult to understand and can you suggest any improvements? Do the authors discuss harms as well as benefits? Are the effects of the treatments over- or understated? **Would be helped by more subheadings after 'Summary of results'**

7. AUTHORS' CONCLUSIONS

Implications for practice: Is this section clear and reasonably easy to understand? **Very valid**
Implications for research: Do you think the authors have identified the important areas for future research? Are there any missing? Are there any benefits or harms important for healthcare users that are not addressed in the studies that you would like to see highlighted here? **Also very valid and comprehensive**

8. DECLARATIONS OF INTEREST

Does the Cochrane Review acknowledge possible interests (e.g. personal or financial) that could have influenced the review authors? **Very clearly**

9. 'SUMMARY OF FINDINGS' TABLE

Are the most important outcomes to you listed in the 'Summary of findings' table? If not, please list them here. **Yes, but, Very limited data for children – lacking in reports…**

10. LANGUAGE AND STYLE OF WRITING ★

a) Is the Cochrane Review reasonably easy to understand? Is the language used clear and well-written? If not, which sections of the Cochrane Review need to be clearer and can you suggest improvements? Is any

language insensitive to consumers? Please suggest alternative phrases if possible.

Is long but all needs to be said, extra headings would help

b) Please list below any words in the Cochrane Review that you think need further definition.

Good explanations of important elements used in review

11. ADDITIONAL COMMENTS

Please add any other comments that you may have:

Well done in this approach and trying to assemble the trial program.
Clearly identifies gaps in the data identified against the defined outcomes. Drug resistance was not covered.
I am assuming all trials were funded by the companies.
Good to have all the background in the appendices.

12. CONFLICTS OF INTEREST

Do you have any potential conflict of interest?	☐ Yes (details below)	**X No conflict of interest**

13. YOUR ACKNOWLEDGEMENT

	Yes	No
I am willing to be identified to the review team as the person who gave these comments.	X	☐
I am happy to be acknowledged in the published Cochrane Review.	X	☐
I am happy to be acknowledged on the Cochrane ARI Review Group website	X	☐

External peer review comments

Cochrane Intervention Review: external peer referee checklist for reviews

A159 – Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children
Referee: [redacted]

ABSTRACT AND PLAIN LANGUAGE SUMMARY

- Do the abstract and the plain language summary accurately reflect the findings and conclusions of the Cochrane Review?

Comment: Yes. However, the objective states the authors will review all clinical trials, both published and unpublished. It is my understanding that published reports were omitted from the analysis due to potential bias. What is meant by "published reports"? Published reports (at least those found in journals) were mentioned in the review as "report biased" and were omitted from the analysis.

"The mechanism of action proposed by the producers (influenza virus-specific), does not fit the evidence." The influenza virus specific action in vivo does come into question; however, it has been shown in a multitude of studies in vitro that its primary action is NA inhibition. Perhaps rephrase this statement to the "sole mechanism of action."

BACKGROUND, OBJECTIVES, AND METHODS

These sections have been previously published in the protocol of this review (available on www.thecochranelibrary.com). However, if you would like to comment on these sections or on any divergence from the protocol, do so here.

Comment: None

RESULTS

- Is there an adequate description of the included studies? Do you get a clear idea not only of what the intervention is, but where it was delivered, when, and by whom?
- Do you have any concerns about how the data has been described or analysed?
- Is there an appropriate analysis of the possible risks of bias in the included studies?

Comment: The description of the included studies is adequate and analysis of the risks of bias included. I have mild concerns regarding the exclusion of studies. I understand that the goal of the review is to provide a clear and unbiased review of complete trial data. 123 studies were omitted for various reasons. The rationale for excluding them is valid but would the conclusions of the review be the same had some or all of the 123 studies (PK/PD excluded) been included in the analysis?

DISCUSSION

- Does the discussion provide an appropriate summary of the results? Do you have any concerns about the authors' interpretation of the results?
- Are the findings set in the appropriate clinical or policy context?
- Does the discussion provide adequate detail about the completeness and applicability of evidence, with specific reference to the quality of the evidence and any potential bias?
- Does the discussion state how the findings of this review compare with other published evidence?

Comment: The summary of the results is appropriate, in clinical context and an accurate representation of the data presented. The criterion of completeness for clinical trial study inclusion into the analysis is an indication

of the quality of the evidence analysed in the review. The authors stated the shortfalls of the clinical trials being protocols and statistical analysis. Were the reports riddled with protocol amendments, study notes and deviations?

The author's state a limitation being inexperience which is only detrimental to the analysis if key pieces of information are over looked or the incorrect analysis is performed. Neither of which would be a shortcoming of this review.

The discussion states that the findings of "alleviation of symptoms" agree with those of other studies. It is the harms or overstated effectiveness that contrasts what was reported previously.

The description of the ferret and mouse study seems out place. I would transition or rationalize its inclusion as to why it is pertinent to human clinical trials. Otherwise omit this paragraph.

CONCLUSIONS

- **Implications for practice:** Are consistent with, and supported, by the results? Can you think of any others?
- **Implications for research:** are they reasonable? Are they specific enough to be helpful in the design, prioritisation, or commissioning of research? Can you think of any others?

Comment: Implications for practice and research are reasonable and supported by the results. I am surprised none of the included trials monitored for NI resistance. At the time of review I could not think of any additional implications already mentioned. Nicely stated!

SUMMARY OF FINDINGS TABLE (IF APPLICABLE)

- Does the Summary of Findings table provide a helpful and consistent reflection of the review and make the key issues clear?
- Did the Summary of Findings table help you to understand the review?

Comment: The Summary of Findings table reflects what is in the review. The table was very helpful when trying to make sense of the findings described in the text.

Comment [CJC1]: Does it add anything to show two lines of control event rates in the SoF? It would be simpler for readers to delete the "Moderate" lines as they are often very similar to the study line anyway. Technically the 95% CI should be (NNTH 16 to infinity and NNTB 48 to infinity) if the 95% CI of the RR crosses one.

ADDITIONAL/GENERAL COMMENTS

- Does the Cochrane Review read well and make sense overall?
- Did you get a clear idea of what the review actually shows regarding intervention effectiveness and any harms?

Comment: The review reads very well and presents compelling evidence to the fact that NIs are an effective influenza intervention in the right circumstances; however, there are "harms" that are often overlooked as a result of the NI medication. Given the frequency of use, we need not only worry about the generation of NI resistant influenza but also the de-emphasized side effects described in the review.

I found the method for collection and analysis to be very interesting approach to getting a full understanding of "real" clinical trial data and how what was reported to FDA for approval. I hope that since the approval of oseltamivir (~15 years ago); the FDA has modified its methods in scrutinizing the clinical trial protocols and data. I have participated in meetings regarding clinical trial protocols for new/novel antivirals and the clinicians refer to running it like was done for oseltamivir. Based on the findings presented in this review, I hope that will not be the case. Once this is published I will refer them to this review. Thank you for all your efforts in putting this together. The authors did a wonderful job.

DECLARATION OF CONFLICTS OF INTEREST

Do you have any potential conflict of interest? ☐ **Yes (details below)** ☒ **No**

PEER REFEREE ANONYMITY AND ACKNOWLEDGEMENT

	Yes	No
I am willing to be identified as the author of this referee feedback	☒	☐
I am happy to be acknowledged in the published review	☒	☐
I am happy to be acknowledged on the Cochrane ARI Group's website	☒	☐

Statistical editor peer review comments

Thank you for the opportunity to comment on this Review. It is a very impressive piece of work. I have only some minor comments on the statistics, plus one major comment

Minor comments

- Page 5 – acronym for intention-to-treat-influenza-infected – should be ITTI rather than ITT
- Page 14 – the statistic "RD = 1.00% (95% CI 0.22 to 1.49, P = 0.05,…)" does not look correct – (I couldn't check from Fig 11 since the results there are given as RR's)

Major comment

- The Review is poorer for the exclusion of the ITTI analysis. There is no doubt that the ITT analysis should be the main analysis presented, and also that there are problems with the ITTI analysis (as detailed by the authors in the Discussion section). However I agree with many of the correspondents (e.g. Frederick Hayden, Helen Steel) that presenting the ITTI information would give important information for clinicians (and it would certainly give better information about the use of Nis in individuals with influenza that is currently available). Although the ITTI analysis could not give 'gold standard' estimates (due to the problems with differential drop-out and the effect of oseltamivir on antibody responses), and of course any such analyses would have to be clearly labeled and cautiously interpreted as perhaps indicating the maximum theoretical potential of Nis, I still think the ITTI analyses would be better off included than excluded

[redacted]

Appendix 7 Anonymised peer review comments and author responses.

Oseltamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. BMJ.2014.017746.R4.

Peer review comments received and responded to March/April 2014

Serial	Comment	Notes	Action
1.	First, we recognize the amount of work involved with this project and we commend you on the excellent work that you have done. The revisions to the section explaining the inclusion process, and describing the evolving methods use in these reviews adds great value as it tells the story to this paper. It will be important to see commonality in this section with the Zanamivir paper – the last version of that paper we saw did not include this detail. We have noted a few errors, and suggestions for improved clarity.	Thank you	No action
2.	Error in abstract results Renal effects when on treatment – the RD value here is –ve when it should be +ve. This difference here appears strange as the RR confidence interval includes one and the RD includes zero. However, the difference is significant using the Peto method in a sensitivity analysis which is reported later on – the Peto method is a more appropriate method to use when events are rare, so this is justified. So the authors are justified here in the strength of their interpretation, although this may be confusing to a reader.	Thank you	Minus sign has been taken out
3.	The inclusion criteria mention post-exposure prophylaxis, but it is not mentioned again in the paper. We presume no studies were found in PEP, but that is not explicitly stated.	Good point; we have added the text on the right	Methods: The Roche trial programme assessing the effects of oseltamivir in post-exposure prophylaxis (PEP) submitted to the FDA on 22 May 2000 consisted of two trials: WV15799 and WV16139. We included only trial WV15799, because WV16139 was not placebo-controlled. WV15799 was a double-blind, cluster-randomised trial in which contact clusters of index cases were randomised to oseltamivir 75 mg a day or placebo for seven days. The manufacturer concluded that the trial proved that oseltamivir could prevent influenza in contacts by interrupting transmission from index cases. Interruption of transmission has two components: reduction of viral spread from index cases (measured by nasal shedding of influenza viruses) and prevention of onset of influenza in contacts measured with a mixture of symptoms and signs and 'laboratory confirmation' (i.e. viral culture from the upper airways and/or at least a four-fold rise in antibody titres measured between baseline and two to three weeks later). The design of the WV15799 is weak. All index cases were left untreated except for a paracetamol rescue pack, making it impossible to assess the effect of oseltamivir on nasal voidance of index cases. Nasal viral voidance was measured only in symptomatic subjects thereby missing out on potential

			asymptomatic infected people. Results Discussion: Similarly to the FDA [18 19] because of the problems with the design of study WV 15799 we could not draw any conclusions on the ability of oseltamivir to interrupt viral transmission. This is important, as the results of trial WV15799 formed part of the WHO rationale for use of the drug to interrupt transmission from person to person and allow time before the arrival of vaccines in the event of a pandemic furnishing a seemingly powerful rationale for stockpiling oseltamivir.[20] This shows the importance of availability of full clinical study reports, something the WHO did not have.
4.	The comment in the methods "We have made a number of changes to the text of A159" line 323 is not going to be understood by readers. What is A159?		A159 has been taken out and "Cochrane review" inserted
5.	The methods lines 421-4 seem bizarre when you read them at this point. They relate to issues considered in detail in the discussion– we question whether they need to be mentioned in the methods at all, and suggest leaving their first mention to the discussion.		The phrase has been deleted
6.	A little more explanation why the results about titre levels etc. (lines 469-78) are harms would be of value		Explanation has been added to the Discussion and text shifted to new line 491 hopefully making it clear
7.	Lines 495-6 would be better placed at the end of the preceding paragraph.		See serial 6
8.	There is a mixture of NNH and NNTH and NNT and NNTB throughout the results. We would suggest standardising to NNTH and NNTB as suggested by Altman. THIS NEEDS TO BE CHECKED IN THE ZANAMIVIR REVIEW TOO.		These have been edited
9.	The first summary of findings tables (Table 5) still includes mean differences in the relative effect column for time to first alleviation for adults and for children. THIS NEEDS TO BE CHECKED IN THE ZANAMIVIR REVIEW TOO.	Thanks you for spotting this.	We have removed the numbers and replaced with N/A
10.	Line 249 states that the RD was computed at the mean control group event rate. However, the summary of findings tables (Tables 5 and 6) state that the median control group event rate was used. Please make these statements correct and consistent. THIS NEEDS TO BE SORTED FOR THE ZANAMIVIR REVIEW TOO.	Thank you for noting this discrepancy.	We have changed the footnote section in Tables 5 and 6 to: "*To estimate treatment effects we first calculated the risk ratios and used the average (mean) control event rate and the pooled risk ratios reported in the figures to calculate the risk differences."*
11.	Other spontaneous edits		NIHR grant link has been updated and reference to Cochrane review inserted. References have renumbered

Comments from BMJ statistical reviewer received and responded to March 2014

Serial	Comment	Oseltamivir Notes	Action
1.	As with the other paper, I would like you to address the additional matters of how you chose the outcomes considered in this study, problems with the definition of pneumonia as mentioned in the other decision letter	The text says: "….of clinical interest into primary and secondary by indication…." We carried out a meta-regression looking at the effects of different data capture methods of the "pneumonia" outcome	
2.	and include a full discussion of the potential limitations to this review along with cautions about overinterpretation of the findings.	Strengths and weaknesses of the review is a whole paragraph in Discussion	
3.	You will find the statistician's comments at the end of this letter. In addition to [redacted] remarks, however, we would like you to take [redacted] comments on the other paper fully into account in this paper as well, since they have many matters in common. I have copied them below -- please recall they were written with regard to the zanamavir paper so that some specific comments will not apply to this paper; you may discount those. Otherwise, please revise this one in line with his suggestions.	[redacted] made no actionable comments: "The statistical analysis of the data is appropriate and the presentation of the results has been restructured to enable a clearer interpretation of the data. Furthermore, a fuller discussion of the limitations of the review is presented. I can find no major statistical issues with this revised paper".	Thank you
4.	In peer reviewing this document I have taken the view that it is of utmost importance to ensure that the science and reporting here is squeaky clean. Hopefully this review is the first of a generation of reviews of evidence freely provided by industry, and it seems to be of utmost importance that it is undertaken to the highest standards, particularly given the criticism that industry has received from the Cochrane Collaboration, these authors, and others, over their reluctance to make trial data publicly accessible.	We agree and thank [redacted]	No action
5.	I have also been provided with access to the full Cochrane Review from David Tovey, editor of the Cochrane Library. My review is quite long and detailed – I hope that I have identified all the points which will allow this review to be improved and made ready for publication. I apologise if I have misread any issues.		No action
6.	Most major points 1) There is no mention of a protocol for this systematic review. I have presumed that it is based on the Cochrane Protocol published in 2011 Issue 1. Including this information in the review is important (PRISMA item 5). There are several ways in which I have noted that the review differs from the protocol: a. Protocol includes comparisons with placebo or standard care, the review only included comparisons with placebo.	Placebo is the only comparator in the parent A159 Cochrane review. [redacted] cited protocol is not the one we followed to do the review. The protocol point is well made	Added to bottom of methods text: **Review Protocol** The review protocol was first published in 2011 (Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of unpublished data. Cochrane Database of Systematic Reviews 2011 ,

			Issue 1 . Art. No.: CD008965. DOI: 10.1002/14651858) and subsequent amendments were published in 2012 and in the current review (see Feedback/Review Amendments 16 May 2013[15])
7.	b. Plans to correspond with the trial's sponsor or report authors where further information is required were included in the protocol but are not mentioned in the review.		Text has been edited to: "**Search strategy** A variety of methods applied to different sources (publications, registries, correspondence with manufacturers, and review of regulatory documents) were used to identify and retrieve manufacturer- and non-manufacturer-funded clinical trials and their clinical study reports"
8.	c. There are no plans for modifying the quality assessment process in the protocol	We used an extended custom-built version of the Cochrane risk of bias tool to appraise clinical study reports. This extraction sheet was finalized prior to but inadvertently not mentioned in our protocol amendments of May 2013.	We have added this to the methods
9.	d. Results will reported in both absolute and relative measures	Comment not understood; we report both	No action
10.	e. There is no mention in the protocol of the analysis by use of relief medications. Some of these points are developed further below.	L339: "finally, data on the effects of rescue or relief medication (mainly paracetamol /acetaminophen) were incompletely reported"	The sentence has been deleted
11.	2) The authors describe a two-step process for selecting reports for inclusion in the review. This is poorly described in the review (a much better description is given in the protocol). Whilst I believe that I understand why this process was adopted (to ensure that the evidence included in the review was internally consistent), the process as described seems subjective and a weak methodology. At first read it appears to be based on assessing whether there is agreement between every document ever written on a study on every detail. I am not sure whether this is the case. Inclusion of a clearer list of the requirements would be helpful. In this review I am not clear whether any trial was actually excluded on this basis. One seems to have been dropped because no CSR was available, and a further one because the comparison was not with placebo (which makes me question why it was included in Stage 1 as it fails to meet one of the inclusion criteria, albeit one which has changed between	In our previous re-write we added: "The main text has been edited as follows: "Because of the novelty and size of clinical study reports we subdivided the extraction, appraisal and analysis of the data into a two-stage exercise. In Stage 1 we assessed the reliability and completeness of the identified trial data. This was particularly important in the early stages of the review when we had received incomplete clinical study reports and were unsure of the importance of the missing parts….". The abstract has been edited to: "We included 23 in Stage 1 (reliability and completeness screen) and 20 in Stage 2 (formal analysis)"	We do not think this comment is applicable to this review

	protocol and review).		
12.	3) In the quality assessment judgements that were unclear because information was not reported were coded as being at high risk of bias. The Cochrane Risk of Bias tool indicates that these should be reported as unclear and not high risk. This change in the way the tool is being used was not specified in the protocol. As a reader I am very interested in knowing the difference between issues which are unclear, and issues which are clearly wrong. The importance of differentiating between poor reporting and poor method has been recognised for a long time, and it is unfortunate that a view has been taken by the authors that it should not be made in this instance. Some unreported issues are unlikely to be sources of bias - for example, the method by which the random order of allocations appears not to be described in the clinical study reports, but it is almost certain to have been done appropriately (as these trials would have been subject to many FDA and MHRA inspections and the thought that they would be done with alternate or other flawed randomisation methods frankly is unbelievable). I am thus really interested to know whether the CSRs did not report this issue, or whether there was evidence that it was actually done wrong. The text suggests that most of the red blobs in the quality assessment arise because of poor reporting - but I cannot tell this from what is reported in the review. It is also a shame that the authors have not contacted the company (the sponsor) for clarification as promised in their protocol.	Comparison with publications is outside the aims of the review	
13.	4) The reporting of data on children is misleading as it largely focuses on statistical significance without reporting estimated effect sizes with confidence limits. The abstract and discussion simply say that the effect was not significant. In the text the treatment effect estimate is given as well. As the point estimate is actually greater than that for adults, simple reporting on the basis of P>0.05 is giving a partial and misleading impression of the findings to a reader. There is no suggestion here that the effect in children is any less than that in adults, which is not what is implied by simply stating the effect in children is not significant. The emphasis on reporting differences simply as “significant or non-significant” without stating effect estimates recurs elsewhere in the review (e.g. interpretation of reduction in asymptomatic influenza) and should be checked through.	This does not seem to be relevant to this review	
14.	5) The results for binary outcomes are presented in the abstract as risk differences and numbers needed to treat. The primary statistic used for meta-analysis was the risk ratio which is not mentioned in	All the required data are reported in the Tables	Added in mean to the paper Added to oseltamivir review

	the abstract. The risk differences and NNTs have been obtained by applying the risk ratio to the "average" placebo group event rate (not stated whether this is a mean or a median, or whether weighting was used in its calculation). The results section states the risk ratios as well as the RDs and NNTs, but does not state the prevalence figure at which RDs and NNTS are computed. The discussion states that influenza was only reduced by a small amount, quoting the NNT, and then questions whether the marketing authorisation was justified based on this.		The following text has been added: "Relative risks and risk differences were used to estimate treatment effects for binary data and mean differences for time to first alleviation of symptoms. To estimate treatment effects we first calculated the risk ratios (RRs) and used the average (mean) control event rate and the pooled RRs reported in the figures to calculate the risk differences (RD). For consistency we adopted this method for both the summary of finding tables and for the RDs reported in the text. For the analysis we chose to report the RRs as they are more consistent across the studies, and we have reported the heterogeneity for the pooled RR"
15.	Whilst I agree with the authors that it is most important to present absolute effects, the framing in this instance is unfortunate. Relative risks tend to make effects look large, whilst in a prevention scenario risk differences will always be small if few people in the study develop the disease, even if a treatment was 100% successful.		No change
16.	Let's look closely at the results reported in Figure 7 and on page 7 lines 27-29. The text states "Zanamivir significantly reduced the risk of symptomatic influenza in individuals, RR = 0.39 (95% CI, 0.22 to 0.70, I2 = 45%); RD = 1.98%, 95% CI: 0.98 to 2.54, NNT = 51 (40 to 103) (Figure 7)". The RD and NNT figures are profiled in the abstract, and the NNT figure in the main paragraph in the discussion (page 9, lines 25-34).		The comment is not applicable to this review
17.	The RR of 0.39 indicates that 61% (nearly two thirds) of influenza cases are prevented by using the drug. In the placebo groups in the 4 trials included here 86 influenza cases were observed in 2644, a rate of 3.26%. Applying the relative risk reduction to this figure would predict that 1.27% would develop influenza if they were taking the drug. The difference between these figures is 1.99% the risk difference, and the inverse of this gives the NNT of 50. Thus the main reason why the risk difference is small is because few people in these studies developed influenza when taking placebo, not because the drug is useful. I believe that it would be most useful for all these figures (i.e. 61% of cases prevented, reducing event rates from 3.26% to 1.27%, giving an NNT of 50 – but all with confidence intervals) to		The comment is not applicable to oseltamivir review

	be reported so that a reader can understand that zamamivir does prevent influenza, but in a prophylaxis situation few people develop influenza regardless of whether they take the zanamivir, and hence absolute benefits are low. Notably the difference in the household analysis mainly occurs because influenza is more common (78 cases from 410 is a rate of 19%, nearly six times as high) and not because the effect of the drug is greater. The authors might want to reflect on the wording of their conclusion statement (Page 9 lines 25-34) to make these points more clearly.		
18.	6) An important analysis reported in the results is that of the impact of zanamivir when given with relief medications. This analysis was not mentioned in the protocol, and I am very confused as to how it has been undertaken. The nature of relief medications is not explained. The abstract mentions "the effect of zanamivir was attenuated by symptom relief medications" and gives figures which it implies are comparing "placebo with relief vs zanamivir without relief". There is no explanation in the methods section contains of these analyses or how the comparison was constructed. The results section mentions that there were seven trials available that allow the comparison reported in the abstract – the data is reported in Figure 5. It is not at all clear to me how these data were obtained and how the use of relief medication was determined. For example, taking the results for study 3008, the characteristics of included studies table (Table 1) states this study had 262 on zanamivir and 263 on placebo. Table 4 reports median values for all participants and participants who did not use relief medication (but does not state how many are in this latter group). Figure 5 gives the total sample sizes but implies that it is using the mean value from the zanamivir group who did not use relief compared and the mean value from the placebo group who did use relief. But everybody in the trial appears to be included in this analysis. There is no division into mutually exclusive groups of those who used relief medication and those that didn't. This can't be right.		The comment is not applicable to the oseltamivir review The comment is not applicable to oseltamivir review. See also serial 8 (6). No analysis of the effect of relief medication was attempted because the data were inconsistently reported across the CSRs We have added to the zanamivir methods A post hoc analysis was undertaken after we discovered 7 trials provided data on time to first alleviation of symptoms with and without relief medication. Each patient in the studies may or may not have taken relief medication during the trial. Alleviation of symptoms may have occurred while the patient was taking relief medication and the "standard" comparison was made using this scenario. However, an additional analysis used a stricter definition where alleviation of symptoms could only be achieved without the use of relief medication. For example, a patient may have achieved alleviation using relief medication after 5 days but took 7 days to achieve alleviation without the use of relief medication. The comparison we report is of all patients where we used the stricter definition for the zanamivir group (alleviation without relief medication) and the less strict

			definition for the placebo group (alleviation with relief medication).
19.	In the discussion the authors state (Page 9 line 4) that "symptoms may be prolonged in the treatment arm when compared to the placebo group on relief medication". It is completely unclear to me where the data to support this statement comes from. The same argument continues in the Implications for practice and research (Page 11 lines 20-22).		The comment is not applicable to oseltamivir Changed zanamivir paper discussion to "However, further analyses reveal the effect upon symptoms is synergised by the use of relief medication, revealing symptoms may be no better in the treatment arm when compared to the placebo group on relief medication." Removed the word "prolonged"
20.	And isn't this analysis an observational rather than subgroup comparison? And is the use of relief medication determined by the participant in response to how they are feeling? There are multiple ways in which such a comparison could be severely biased (including being driven in the zanamivir group by response to the drug) which need to be acknowledged. I am not convinced that the authors' conclusions for these analyses are justified (and certainly the analysis as presented is erroneous).		The comment is not applicable to oseltamivir review
21.	Less major points 7) The phrase "clinical study report" is not widely understood (I have tested the phrase on several colleagues who did not know what was meant by it). The authors should find an alternative phrase or give a better explanation.	Clinical study report is an official ICH term. In the last 5 years the term and content has received extensive coverage. The text in the Introduction is supported by reference 9 which is an open access exploratory review of CSRs of 14 different drugs. Because of space restrictions I am not sure we can do more.	We have inserted further explanation in the introduction: "In the case of oseltamivir clinical study reports mean length is approximately 1305 pages (median around 900 pages)".
22.	8) I would have liked to read a clearer description of the search strategy in the text. Also, the paragraph explaining the inclusion criteria (Page 2 line 56 to Page 3 line 18) would be easier to follow if it organised the criteria according to the PICO elements plus description of the study design, as in the Cochrane Protocol and Review. It is not clear what the implications from the comment about "pivotal trials" (Page 3 line 2) – did these not have clinical study reports?	The relevant text says: "There was a mix-up with follow-up cards in the "pivotal" trials WV15670, WV15671 and WV 15730 which does not allow drawing any conclusions on the durability of symptom relief[21]".. and "Information on a problem with follow up cards in three "pivotal" treatment trials was only discovered thanks to FDA SBA papers". The implications are well described	No action
23.	9) The report refers three times to Appendix 2, each time	Appendix 2 (Searches for Regulatory information) was	The comment is not applicable to

	seemingly to a different Appendix 2, none of which are provided.	provided	oseltamivir review
24.	10) The extended quality assessment list (bullet point list reported on page 6 lines 37-50) is also far from standard assessment, and it is not clear what the value of this information is in assessing risk of bias. This reads more like the list of findings of an MHRA CTU inspection visit looking at execution according to Good Clinical Practice guidelines and Standard Operating Procedures than assessing items which are known to link to bias. These items were not specified in the protocol.	This comment has to be inserted in the context of the time it took us to get the complete set of clinical study reports (4 years). In 2011 we asked Roche a series a clarification questions which were not answered. We have carried out changes to the risk of bias tool and our methods of using it but they require a self-contained paper to report these.	Previous reviewers asked for more information
25.	11) The quality assessment criteria reported in Figure 2 are not described. I am particularly baffled by the "other bias" category. Some explanation of the domains in the methods or as a footnote to the figure would be welcome.	The BMJ Higgins paper categorises other bias as any other bias identified by the researchers which does not fit in the other categories. [redacted] is thinking publications, we dealt with clinical study reports. For example the presence of dehydrocholic acid in the placebo and the different coloured placebo cap are examples of types of other bias which will only be detected in Clinical study reports	For explanation purposes we have added the text underlined The placebo capsules in oseltamivir trials contained dehydrocholic acid and dibasic calcium phosphate dehydrate (we classified these as other potential biases).
26.	12) The conclusion in the abstract and text does not mention the effect of zanamivir on prevention at all. This seems an oversight.		The comment is not applicable to this review
27.	13) The authors reports a comparison of treatment effects by infection status (Page 7 lines 18-22 and Figure 6) reporting a test for difference in subgroups. This nature of this test is not reported in statistical methods section. In contrast to the analysis by use of relief medications, this analysis does divide the participants into mutually exclusive groups according to their infection status.		Comments not understood
28.	14) The number of figures reported is excessive (18), and could easily be reduced by combining several plots into single figures. For example, the impact on different definitions of influenza and pneumonia could be put on single slides, similarly the harms.		No change We are happy if editors want to redraw the figures and consolidate them
29.	15) I would have liked to have a clear comparison of the results of this review based on the hard-won CSRs with that of previous reviews which used only data in the public domain. The authors state that not using all available data introduces bias but they do not tell us how		

	it has affected the findings of this review.		
30.	16) Page 9 line 37 - the BMJ has previously published a systematic review based on all clinical study reports supplied by industry – http://www.bmj.com/content/325/7365/619.	This indicates a more widespread understanding of the term CSR, thank you	

BMJ peer review comments received and responded to February 2014

Serial	Editor/Referee point Only critical comments are reported, not agreeing comments or those not requiring/asking for action	Comments	Action
1.	* As we mentioned in the decision letter for the companion paper, these papers are very important because they are going to be pored over and they may lead to legal action. We therefore feel they must be complete and live up to standards we expect for other systematic reviews and meta-analyses. Otherwise we worry that, having criticized industry for not providing detail you will be criticized for not providing it yourselves. We felt unable to appraise the statistical aspects of the paper because the data were not there -- all we had was the summary of results tables.	We tried to make things simple, it did not work. The story of the review is complex and some of the methods reflect this. A good example of this is Stages 1 and 2 of inclusion issue (see serials 34, 41 and 42). We devised the stages as a means to cautiously guide us in the use of the EMA clinical study reports which were A. new material to us and B. all (except for one) incomplete. This is the reality, not the theory, but we think that its detailed description is going to make the text very lengthy	The text has been re-written, lengthened and much more information has been provided in the hope of addressing this comment. The results have all been updated and associated figures are included
2.	Please return the paper with a full accounting of the methods and Forest plots for the main results so that we can more fully appraise this.		We are including slides of 28 forest plots. Perhaps all the slides could be included in the online publication.
3.	* Please note that although one of the reviewers suggests you might editorialize a bit in this paper, we do not agree -- those things can go in the accompanying Analysis article and will also be points for an editorialist to make.	We agree, that is why we wrote an Analysis	The comment will be ignored
4.	* Table 3 – it is hard to interpret number of events in isolation from number exposed; can you indicate this somewhere?	We agree and have added the total number of patients included in each treatment arm	Table 3 has been revised (now Table 4)
5.	* We weren't clear about how decisions were made to exclude studies if they did not have "consort statement specified results." We worry that many of the decisions made during the course of this review are not clearly described and wonder if anyone would be able to replicate this.		The relevant text has been edited. We hope it is more comprehensible
6.	IMPORTANT When you revise and return your manuscript, please take note of all the following points. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision.	Understood	No action
7.	a. In your response to the reviewers and committee please provide, point by point, your replies to the comments made by the reviewers and the editors, and please explain how you have dealt with them in the paper. It may not be possible to respond in detail to all these points in the paper itself, so please do so in the box provided		This table represents our response
8.	b. If your article is accepted it will then be edited, proofed, and - after your approval - published on bmj.com with open access. This open access Online First article will not be a pre-print. It will	Understood	No action

	represent the full, citable, publication of that article. The citation will be year, volume, elocator (a unique identifier for that article): eg BMJ 2008;337:a145 — and this is what will appear immediately in Medline, PubMed, and other bibliographical indexes. We will give this citation in print and online, and you will need to use it when you cite your article.		
9.	c. Please write an abridged version of the article for the print and iPad BMJ using the appropriate BMJ pico template for your study's design. Please be reassured that it doesn't take long to complete this. When your BMJ pico is ready please email it to papersadmin@bmjgroup.com. The templates for you to download are at http://resources.bmj.com/bmj/authors/bmj-pico		The required text is included in the manuscript
10.	Please include the items below in the revised manuscript to comply with BMJ style: * the title of the article should include the study design eg "a retrospective analysis of hospital episode statistics"		The required text is included in the manuscript
11.	* ID of ethics committee approval and name of the ethics committee/IRB; or a statement that approval was not required (see ttp://resources.bmj.com/bmj/authors/editorial-policies/guidelines)		The required text is included in the manuscript
12.	* Please complete the following statement and insert it in your manuscript: Competing interests. Please complete the following statement and add it to your manuscript: "Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work [or describe if any]; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years [or describe if any]; no other relationships or activities that could appear to have influenced the submitted work [or describe if any]." (Please see http://bit.ly/T7uoG2)		The required text is included in the manuscript
13.	* contributorship statement+ guarantor (see http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship)		The required text is included in the manuscript
14.	* copyright statement/ licence for publication (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)		The required text is included in the manuscript
15.	* signed patient consent form(s), if the article gives enough personal information about any patient(s): this sometimes occurs even in research papers - for example in a table giving demographic		The required text is included in the manuscript *Ethics approval and patient consent forms are not provided as they are*

	and clinical information about a small subgroup in a trial or observational study, or in quotes/tables in a qualitative study –		*not necessary for a Cochrane review.*
16.	http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/patient-confidentiality * for a clinical trial, the trial registration number and name of register – in the last line of the structured abstract		Not applicable
17.	* for any other registered study (eg a systematic review), the registration number and name of register – in the last line of the structured abstract		Not applicable to a Cochrane review
18.	*a data sharing statement declaring what further information and data your are willing to make available. Suggested wording: "Data sharing: technical appendix, statistical code, and dataset are available at this repository or website <url> OR from the corresponding author at <email address or url>". If there are no such further data available, please use this wording: "Data sharing: no additional data available"	We will make all CSRs for the two drugs available. Discussions are underway with Editor in chief as to where	The statement "Data sharing: all clinical study reports are available at this repository <url> " has been inserted
19.	* please write the discussion section of your paper in a structured way, to minimise the risk of careful explanation giving way to polemic. Please follow this structure: * statement of principal findings of the study * strengths and weaknesses of the study * strengths and weaknesses in relation to other studies, discussing important differences in results and what your study adds. Whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses (eg Cochrane reviews) * meaning of the study: possible explanations and implications for clinicians and policymakers and other researchers; how your study could promote better decisions * unanswered questions and future research		We have restructured the discussion as you suggest. The two main limitations that we could spot (inexperience and applicability outside clinical study reports) have been added to the Discussion.
20.	* please note, too, that the article's introduction should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now	See serial 32	
21.	* What this paper adds/what is already known box (as described at http://resources.bmj.com/bmj/authors/types-of-article/research)		
22.	* funding statement (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements)		The required text is included in the manuscript
23.	* statement of the independence of researchers from funders (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements)		The required text is included in the manuscript

24.	* for studies funded or sponsored by industry (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements):		Not applicable
25.	* a statement describing the role of the study sponsor(s), if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication		The required text is included in the manuscript
26.	* assurance, in the cover letter, that a clinical trial funded by a pharmaceutical or other commercial company follows the guidelines on good publication practice (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements) * inclusion in the list of contributors the name(s) any professional medical writer(s), specifying in the formal funding statement for the article who paid the writer. Writers and authors must have access to relevant data while writing articles. * structured abstract (see http://resources.bmj.com/bmj/authors/types-of-article/research)		Trial funding issue is not applicable. No medical writer was involved, only us! Structured abstract is included in the manuscript
27.	* summary statistics to clarify your message We do want your piece to be easy to read, but also want it to be as scientifically accurate as possible. Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate: For a clinical trial: • Absolute event rates among experimental and control groups • RRR (relative risk reduction) • NNT or NNH (number needed to treat or harm) and its 95% confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000) For a cohort study: • Absolute event rates over time (eg 10 years) among exposed and non-exposed groups • RRR (relative risk reduction) For a case control study: • OR (odds ratio) for strength of association between exposure and outcome For a study of a diagnostic test: • Sensitivity and specificity • PPV and NPV (positive and negative predictive values)		Understood
28.	For research articles As well as submitting your revised manuscript, we also require a		This will be done

	copy of the manuscript with changes highlighted. Please upload this file with file designation 'Revised Manuscript Marked copy'.		
29.	REFEREES COMMENTS Reviewer: 1 This manuscript describes an important investigation with wide implication for: 1) prophylactic and therapeutic treatment of influenza; 2) standard methodology of systematic reviews of drug/device interventions, and 3) the lack of access to primary study data in drug/device trials.	Thank you	No action
30.	I would like to congratulate the authors on their pioneering task of identifying, assessing, and summarising the unpublished data on oseltamivir for influenza, and thus to present a sketch for a road map for how to deal with the problem of unpublished trial data in other clinical settings.	Thank you	No action
31.	The political, clinical, and methodological ramifications of this investigation are important so the manuscript should live up to quite high standards for methodological rigour and for clarity of communication. I have, with this in mind, some concerns with the manuscript, mainly to do with presentation and style, but also to do with methodological issues: 1) The presentation is generally unclear. It is challenging to re-write the text for a Cochrane review (rich on detail and methodological aspects) so that it fits the more compact style of the BMJ (with focus on general readability and clarity). In my opinion the text could be improved considerable. For example, I suggest that: the authors avoid the "evidence jargon" (for example to avoid phrases like "the evidence shows ..."), explain what is "NI", and explain that Roche is the only (?) producer of the drugs, avoid emphasis on "statistical significant" (instead of the point estimate).	We tried to be as clear and factual as possible, leaving our more private thoughts to the Analysis piece but we obviously failed.	We agree with the remark on the evidence jargon and have edited the text accordingly throughout
32.	2) The introduction is very short, and provides too little background for understanding the implication of this study.	See serial 20 your instructions. Even with 3 paragraphs it is impossible to cover the last 4.5 years's work, which is why we wrote the analysis	A 3rd paragraph has been inserted with outline detail on significance and content of clinical study reports
33.	The Discussion is not structured according to the standard BMJ disposition. I miss a clear reflection of strengths and weaknesses of this study, relation to other similar studies, and implications.		See serial 19
34.	The abstract contains very little information on methods, and is not self-explanatory (what is stage 1 and stage 2), manufactures are mentioned first and then Roche, the result section is very large and difficult to read.		The main text has been edited as follows: "Because of the novelty and size of clinical study reports we subdivided the extraction, appraisal and analysis of the data into a two-stage exercise. In Stage 1 we assessed the

			reliability and completeness of the identified trial data. This was particularly important in the early stages of the review when we had received incomplete clinical study reports and were unsure of the importance of the missing parts….". The abstract has been edited to: "We included 23 in Stage 1 (reliability and completeness screen) and 20 in Stage 2 (formal analysis)".
35.	The PICO template includes the phrase "regulatory evidence" without explanation?		The text in red has been added to the introduction: **Selection criteria for studies**: Clinical study reports of RCTs testing the effects of oseltamivir for prophylaxis and treatment of influenza in healthy people or the chronically ill who have symptoms of influenza-like illness. These were augmented by regulators' comments and reports during drug registration.
36.	The issue of inexperience of dealing with large datasets from trial reports is mentioned in the PICO but not in the Discussion of the paper.		See serial 19
37.	In the central tables 4 and 5, why is the degree of heterogeneity for each analysis not reported?		We have now included forest plot figures for all the analyses reported; and these include the degree of heterogeneity
38.	The first of five main outcomes (listed in the abstract) is time to first alleviation of symptoms. No description of primary and secondary outcomes or criteria for selection of outcome is presented in the text, or planned (and unplanned) sub-group and sensitivity analyses. I suggest this is added.		We have now added the requested hierarchy of outcomes
39.	Time to event outcomes are ideally and usually summarised as hazard ratios or as ratio of medians. Mean time to event will normally suffer from the problem of how to deal with censoring, and even with a full dataset, the data will often be skewed. Only in a scenario with full data and no censoring (all patients have the event) will mean time be easy to interpret. One trial has a follow up time of 6 days, with high risk of censoring. I suggest that the authors comment on why time-to-event data were treated as if measurement scales outcomes, and whether this implies a study limitation.	We used hazard ratios in our previous reviews in 2006 and 2009 but were criticised by GSK as they said the proportional hazards assumption is not met for this outcome in their trials. Also reporting treatment effects in terms of hazard ratios is problematic. The alternative in RevMan is to use means and standard deviations as we have done. We agree there is a limitation with patients who did not reach the endpoint leading to under-estimates of means. However the proportions of censored patients was low in all trials and very low in most trials. In addition the proportions censored were similar in both treatment groups hence we do not believe there is a bias. As you	We have explained our reasoning for using means and standard deviations and also included this as a limitation in the discussion.

		say, another alternative is to use medians but these too have limitations as they only represent the middle of the distribution.	
40.	The selection of outcome is not easy to follow. Were they pre-specified? How were they handled?	The methods text has been re-written. The outcomes of interest were pre-specified in our protocol but the yardstick was of interest to clinicians in our review group	The text now reads: "We divided the outcomes of clinical interest into primary and secondary by indication as follows….."
41.	The risk of overlooking important information or data in 150.000 pages is considerable. What was done to prevent missed data, to check whether data had been missed, and to ensure that data extraction was not biased?	Good comment. Quantitatively precision of estimates are reported as 95% CIs and heterogeneity has been investigated where possible in subgroup and sensitivity analysis and use of random effects models.	The following has been added: "We used a two stage process for data extraction and quality assessment via use of customized data forms; these were done independently by two reviewers , while a third reviewer arbitrated. Secondly, the extracted data on quality of the studies was again corroborated by a meeting of all authors. Thirdly, the data on outcomes for statistical analysis was independently cross-checked to ensure that numbers presented in the forest plots matched actual data from the clinical study reports".
42.	The selection of trials appears difficult to follow. Under data extraction it seems as if trials were included/excluded not on the basis of whether they had been conducted, but on the basis of how they were reported?		The text has been edited to clarify, see also serials 1, 31 and 41
43.	The review could more clearly differentiate between inclusion in the review (n=23) and inclusion in the meta-analyses (n=20). I suggest that a short paragraph is added describing the three trials with incomplete clinical study reports.	See serial 84	
44.	The strengths of the study is clearly stated, but I miss a reflection on: 1) the risk of misrepresenting and missing data due to the sheer size of the dataset, 2) risk of bias in the trials and other aspects of "quality of evidence" (e.g. heterogeneity, precision of estimates); 3) using mean time instead of median time (or hazard ratios) to summarize effect (see point 8).	See serials 39 and 41	
45.	I also think it would be interesting to have the authors reflected opinion on the balance of possible benefits and possible harms of the intervention, and their suggestion for the implication of their study (clinical, political, scientific).	The political implications will be tackled by Editorial and Feature	
46.	Reviewer: 2, Comments: Major issue This is a terribly important paper that I am convinced the BMJ will publish. But I have a major reservation that I feel needs to be		We have added "Our findings do not support the stockpiling of oseltamivir, its inclusion in the WHO's list of essential drugs nor its use in

	addressed: No people in the whole world know more about Tamiflu than the authors of this systematic review. I therefore feel they need to be much more direct in their conclusion and paper. As far as I can see, there are so many uncertainties, so much bias, and still the effect, given all that, is so small or not discernible, if people have used paracetamol (Glaxo trials), that I feel the authors should suggest not to use Tamiflu at all. It seems to me to be a classic example of how much the industry can cheat us with their randomized trials.		clinical practice as anti-influenza drug". Considering its toxicity its minimal benefits and the likely mode of action, an anti-influenza role is meaningless
47.	Furthermore, I lack a mentioning of the discrepancy between the FDA and the EMA, where the EMA accepted Roche's completely unfounded claims about reduced complications and hospitalizations, which the FDA didn't. The authors should not simply reproduce a condensed version of their Cochrane review but should use the Cochrane review to tell the world some very important lessons. Tom Jefferson uses to say that he has been 'Kaiserized', alluding to the terribly flawed meta-analysis Kaiser and his co-author did on unpublished Roche trials, financed by Roche, that claimed fantastic effects. Jefferson et al. struggled for years to see these unpublished data, and based on their now updated Cochrane review, my conclusion is that Roche has committed serious fraud. The BMJ lawyers may not allow such an accusation, but please do put your findings into perspective and leave no doubt about how Roche has cheated the whole world with their Tamiflu trials. Rather than praising the company (see below), which I fail to understand the authors do after what Roche has put them through.	The reviewer cites some important evidence on the multi system failure of the Tamiflu story that we first published in 2009. We mention the issue of discordant regulation and who saw what in our Analysis. The reviewer should have access to it.	No action
48.	ABSTRACT There are two sets of line numbers, therefore difficult to know which one to refer to; I have chosen the column closest to the text.	We don't know what the second set of line numbers is	
49.	Don't use terms such as 'no evidence of 'and 'insufficient evidence', as it is not clear what is meant by this.	We agree	The words have been edited out
50.	L 79: Please write by how much. (risk of influenza in px trials)		Abstract text has been edited to: Abstract In prophylaxis trials, oseltamivir reduced the risk of symptomatic influenza symptoms in individuals and households by 55%, (RD = 3.05% (1.83 to 3.88); NNT = 33 (26 to 55), there was no significant effect on asymptomatic influenza. In the results it reads as

			Oseltamivir reduced the risk of influenza symptoms in individuals by 55%, RR = 0.45 (95% CI, 0.30 to 0.67, I2 = 0%); RD = 3.05% (95% CI: 1.83 to 3.88); NNT = 33 (95% CI: 26 to 55) (Figure 16);
51.	L 85: Not possible to have an NNH of >8109 when there was no significant difference (in that case the upper NNH is infinite).		The NNT has been deleted
52.	PICO, L109, Summary answer: "The evidence shows that oseltamivir causes vomiting, nausea, headaches and psychiatric syndromes but may protect recipients from cardiac events while shortening duration of influenza-like illness self-reported symptoms in treatment and preventing their appearance in prophylaxis." The summary answer is not clearly written, and where does suddenly the cardiac events come from and was this an outcome in the trials or the Cochrane review, which I would doubt? What does: 'preventing their appearance in prophylaxis' mean?		The cardiac events line has been deleted. We are unsure to this day whether this is harm or a benefit.
53.	L125: Unclearly written (in adults and in children). The results section should be rewritten and made clearer, e.g. L 143 mentions renal events twice. Main text		The repetition has been deleted
54.	L 181: What is 'open literature'?		The word has been edited to "publications"
55.	The authors should avoid using the word evidence, which is used far too much these days, for example in L 189 'we included evidence from RCTs', why not just we included data?		The word has been edited out throughout where appropriate
56.	L 193: 'Open label studies', this term should never be used, it is misleading, as it usually means a study without a control group, whereas the name implies that it is an unblinded study, which is something else.		The word has been edited to "unblinded"
57.	L 204: Why the word both?		The word has been edited out
58.	L 208: 'broadly consistent', this leaves some room for arbitrariness. The authors should explain carefully, either here or by saying that they did this in the Cochrane review, which studies they excluded and why.		The word "broadly" has been edited out
59.	L 242: It would be interesting to know what incomplete reporting in the clinical study report means, particularly since the authors excluded some trials for lack of internal consistency.	Internal consistency was a criterion for entry to Stage 2 in the earlier versions of the review. This has been clarified – see also serial 41	A list of missing data has now been added
60.	L 251: It is not possible to calculate a NNT for a difference that is not statistically significant and negative NNTs do not exist.	Not strictly true see Altman et al (BMJ) http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1114210/	The NNT however has been deleted

		There is an argument that one does not wish to know the number needed to treat unless there is clear evidence of effectiveness, which for convenience alone is often taken as having achieved P<0.05. This advice seems to be based, at least partly, on trying to avoid the difficulty of an infinite number needed to treat rather than statistical soundness. In fact, we might often wish to quote a confidence interval for the number needed to treat when the confidence interval for the absolute risk reduction includes zero. Though this can be done by quoting two separate intervals, such as NNTB 10 (NNTH 20 to ∞ and NNTB 4 to ∞), I suggest that it is done as, for example, NNTB 10 (NNTH 20 to ∞ to NNTB 4), which emphasises the continuity.	
61.	L 252: Is an example of the meaninglessness of the term 'insufficient evidence'. What does this mean? Too few patients or too events, or what?		See serial 55
62.	L 254: It is very important information that crucial details of trial methodology may only be revealed by having access to case report forms.	Refers to "All oseltamivir treatment trials collected data on complications through participant self-reporting mediated by an investigator who filled out a form"	.
63.	L 259-261: Again, rather than using the evidence word, please tell us what the result showed!		Twenty eight additional Figures have been added to the submission
64.	L 264: Please give us the data.		Figures have been added to the submission
65.	L 270: Don't you mean influenza like illness here and elswhere? Explain clearly what you mean every time.	The text is correct	No change
66.	L 273: Please explain what is meant by asymptomatic influenza and what the data showed.		The following has been added as an explanation: " (i.e evidence of infection such as antibody rises in the absence of symptoms)"
67.	L 277: Using NNT for baseline differences is a contradiction in terms, as no one has been treated. Please explain here under results why you mention these baseline differences and what they mean, as they are surprising given appropriate randomization.	The reviewer's confusion is because of a lack of understanding that the "influenza infected as baseline" subgroup is not actually determined at baseline	The relevant text now reads: "…..compared to the control group (RD = 4.71%; 95% CI: 1.77 to 8.24; P=0.01; I^2 = 21%; NNT = 22, 95% CI:13 to 57). The baseline differences are identifiable only retrospectively, as influenza diagnosis, partly based on antibody response, is made during a 2-4 week follow up. Oseltamivir in treatment….."
68.	L 280: Don't write 'associated with'. The reason we do randomized trials is that we can make causality inferences, 'associated with' is something that comes from epidemiology. Please write that oseltamivir increased the risk of nausea and nothing about		We have edited where possible

	'associated with'. Correct throughout the paper.		
69.	L 284: The decreasing cardiac events were not significant, I therefore wonder why the authors mention this in their PICO, which they should not do.		See serial 52
70.	L 286-88: Sentence not clear. This paragraph deals both with psychiatric adverse events and antibodies. Should be separate paragraphs.		Antibody and psychiatric effects are now listed under “harms” para
71.	L 301: Not possible to calculate NNH here.		The NNT has been deleted
72.	L 313: It is not a family of drugs, it is one drug.		“(on neuraminidase inhibitors, of which this review is part)” has been added
73.	L 316: The authors should not recognize the role of Roche without also saying that Roche was struggling hard for years against releasing these data, and that this would never have happened without the persistence of the authors and that of the BMJ. Roche is certainly not to praise! Roche is to blame for having stolen all this money from the taxpayers for a drug that is no better than paracetamol.		Roche acknowledgment has been deleted. BMJ stays
74.	L 319: 'Against' why use this word?		This has been deleted
75.	L 324: It is not possible to conclude that oseltamivir has anti-inflammatory properties based on the data the authors reviewed, and I assume, based on any data available on oseltamivir. The authors need to discuss here that the slight reduction in duration of illness could easily be caused by bias, since it is very subjective to decide when an influenza-like illness is over. Hróbjartsson has shown in BMJ in 2012 that unblinded observers exaggerate subjective outcomes by 36% on average measured as odds ratio compared to blinded observers. The whole effect could therefore easily be caused by unblinding due to the conspicuous side effects of the drug. The authors should quote Hróbjartsson’s important study, I believe.		The text has been edited accordingly
76.	L 327: What is meant here by cardiac rhythm?		“ECG anomalies” inserted
77.	L 334: Please expand on the explanation about antibody responses; it is very clear to you, but not to the readers.		The text has been edited accordingly
78.	L 345: It is very interesting that the drug may have an antipyretic effect, and the authors should mention somewhere that the small effect obtained by oseltamivir could very likely be obtained by paracetamol, for example. Also that the trials of Glaxo’s drug, zanamivir, could not show any effect when the patients also used antipyretics.		The text has been edited accordingly. The zanamivir review has more data on the effects of rescue medication. Roche CSR did not report it for whatever reasons
79.	L 350: Why would an influenza-like illness relapse after 5 days? You need to describe that what you mean is likely the fever, and that therefore the recorded disease durations may be wrong and similar	We have explained again (in the list of missing data and in the Discussion) the potential impact of the mix up with follow up cards. FDA did not know and neither	

	to selling rubber band by the meter, which is rather flexible. What happened? Were disease durations truncated by five days or what (which would be terribly wrong)? It appears to me that one cannot draw any conclusion about the durability of symptom relief, and it therefore surprises me that you give the duration of illness in hours with one decimal. Please explain and leave out the decimal. Can you say anything about the effect of oseltamivir on disease duration given all these problems? I doubt it.	do we as the data were not collected.	
80.	L 356: The readers cannot make anything out of the comment about the drug having no effect on asthmatic children, since we are only given information like 'insufficient evidence'. We haven't seen any data, and there could just be too few children to draw any conclusions, which is not the same as 'no effect'.		We have included figures with forest plots and "evidence" has been deleted throughout where relevant to this comment.
81.	L 372: This sentence is not very clear.		Now: "In the psychiatric events MedDRA System Organ Class (http://www.meddra.org/how-to-use/basics/hierarchy) several rare Preferred or Lowest Level Terms representing single events (nervousness, aggression, suicide ideation, paranoia) reported significantly more frequently in the intervention arm, added to other more frequently reported but not significantly different events (such as depression) gave a large effect and relatively small number needed to harm of 81 (33 to 556)".
82.	L 387: This explanation should be made more easily understandable.	This is not easy without inserting conspiracy theory stuff which we try to avoid. The most baffling fact is that we used Roche's own MeDDRA codes untouched. We simply don't know. Compliharm play is the most likely explanation	.
83.	L 405: What does 'under recruited' mean?		The text has been edited to: "Treatment trials were mostly under recruited (achieved sample size below planned sample size with the explanation that influenza circulation was considerably below expected levels) and often their results were pooled in two or even three trials and yet they showed very high influenza positivity rates (up to 80%)."
84.	Flowchart: 60 studies out of 83 did not meet exclusion criteria. I know that this is detailed in the Cochrane review, but the readers		The text now reads: "Sixty studies were excluded (Figure 1 and web extra), 40 (67%)

	would be interested in why so many studies were thrown out. This appears from page 25, but a summary needs to be given in the main text, particularly since most of these trials did not have a control group or were just pharmacokinetic studies.		of which were pharmacokinetic and 10 (17%) were unblinded and/or non-comparative studies" And "The first two were synopsis translations and the third was 50 page Roche-Shanghai internal report." [this refers to the 3 excluded trials] Serial 43 also refers
	[redacted]		
85.	The tone in the introduction and elsewhere made the authors sound as if they were anticipating the results before starting their review -- in other words the tone could strike some readers as biased. It is imperative that the tone remain neutral in the introduction, methods and results, if only to convince the reader that they have not decided ahead of time what the answer is.	We appreciate the reviewers' comment. But the rationale for including only clinical study reports and complete trial programmes is based on the sizeable reporting bias that we have extensively documented	We are open to suggestions
86.	I was a bit confused by the discussion about "self-reported" pneumonia. On page 9, line 319, the wording has me confused. The drug appears to protect against the "complication" pneumonia when the criteria are less strict (ie, the association is statistically significant), and it is not protective when the stricter criteria are imposed, correct? Yet Table 4 refers to the less strict criteria as "self-reported," and in the next line, contrasts the findings with "clinician-diagnosed" pneumonia. The RR for the two ways of defining pneumonia are about the same, but the 95% CI is narrower for the less strict criteria and does not include 1. This contrasts with the findings for the more strict criteria ("clinician-diagnosed pneumonia"). So the interpretation is that in both cases the drug is protective but this protection was not statistically significant when stricter criteria for diagnosing pneumonia are imposed, right? So why do the authors call the less strict criteria "self-reported" and the more strict criteria "clinician diagnosed"? Aren't the two outcomes measured the same way but with less strict criteria in one case, or is something else happening? Tell us a little bit more, how were the criteria applied? Was it two different forms were used, was there a committee that decided whether the participant had pneumonia, did one group have a lab test and the other did not? Also, isn't it more typical to show all pneumonia diagnoses together and then do a sensitivity analysis of the two categories of strictness separately? What were the overall findings? Anyway, given that this is a possible benefit for the drug, and the review overall is pretty negative about the drug, it seems	Some (mainly early) trials collected secondary illnesses in a similar way to adverse events via self-report from patients on adverse event/intercurrent illness forms. These forms included no information on how the secondary illness was diagnosed. Other trials used a different form that included detailed information on diagnosis including chest X-rays. Based on this observation we pre-specified in our protocol that we would stratify our analysis by method of diagnosis.	We have now included the results of a meta-regression from our Cochrane review on the pneumonia outcome reported in all included studies of oseltamivir and zanamivir that shows method of diagnosis is the most plausible explanation for the heterogeneity of treatment effects we observed across the various treatment populations. We have also included a new table (T 1) showing the data capture in the relevant parts of the case report forms.

	preferable to show any possible benefits clearly and in an up front way so that your findings are not perceived as biased or pre-determined.		
87.	Major structural changes		Tables have been re-numbered in the text. 28 Figures are now submitted with the manuscript

Appendix 8. Anonymised peer review comments and author responses.

February 2014

Dear Dr. Heneghan,

Re: Manuscript ID BMJ.2014.017753 entitled "Zanamivir for preventing and treating influenza in healthy adults and children: systematic review of clinical study reports"

Report from the BMJ's manuscript committee meeting

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Decision: Put points, revision to be sent to [redacted] [[redacted]] for review.

I know that both of the systematic reviews of Zanamavir and Oseltamavir are really important reviews, and that the BMJ will want to publish them, but these reports are poor, and the authors need to undertake substantial work to turn these into scientifically credible docs. A huge amount of detail is missing, both in the methods, and most importantly, the results. They do not present any results from the individual studies. They are a long way from complying with the reporting guidelines here.

I would suggest that they are returned with the peer reviewer comments (as there are so many) and a clear direction to include full methods and the forest plots for the main outcomes, then they undergo a statistics review. There is little point in doing such a review without the study results in the document

As for the accompanying paper, we want to publish these studies, perhaps with an accompanying editorial highlighting why these are different from other MAs and how they are an example of using open data. But I agree with comments above that a lot of work is needed to improve the reporting. It is not enough to summarize the Cochrane paper leaving out information.

Response: We see that this point comes up throughout the peer review comments. In our attempts to be succinct we see that we have overdone it, missing out important details as you state, including sections of the methods and the results, which do appear in the Cochrane review. We have therefore rectified this supplying all of the relevant figures, which was discussed at the joint BMJ/Cochrane phone meeting, on the 13th February, we have added more of the data on harms and the relevant effect sizes.

As per the point above, we have included important information and outlined a point by point response below:

Point number		Comments	Action
1	These authors deserve credit for their tenacity. They seem to be really making a difference, by not giving up, and by admirable attention to detail. [redacted]'s comments are well made though, so despite all that detail, we need still more. For me, this is more about the battle for open data, than the effectiveness of these agents. Here's what can be achieved when people are willing to dedicate their lives to it. The value of neuraminidase inhibitors for the prevention and treatment of seasonal influenza: a systematic review of systematic reviews. Michiels B, "The combination of diagnostic uncertainty, the risk for virus strain resistance, possible side effects and financial cost outweigh the small benefits of oseltamivir or zanamivir for the prophylaxis and treatment of healthy individuals. No relevant benefits of these NIs on complications in at-risk individuals have been established." Its value is mainly historic, I think, as they now were able to analyse many more trials than the 10 they had in their last MA: Cochrane Database Syst Rev. 2012 Jan 18;1:CD008965. doi: 10.1002/14651858.CD008965.pub3. "…25 studies (15 oseltamivir and 10 zanamivir studies). We could not use data from a further 42 studies due to insufficient information or unresolved discrepancies in their data. We found a high risk of publication and reporting biases in the trial programme of oseltamivir. Sub-population analyses of the influenza infected population in the oseltamivir trial programme are not possible because the two arms are non-comparable due to oseltamivir's apparent interference with antibody production. The evidence supports a direct oseltamivir mechanism of action on symptoms but we are unable to draw conclusions about its effect on complications or transmission. We expect full clinical study reports containing study protocol, reporting analysis plan, statistical analysis plan and individual patient data to clarify outstanding issues. These full clinical study reports are at present unavailable to us."	Many thanks We agree with the detail issue, and see that this is a point reiterated throughout the peer review comments	
2	Similar to the other paper: lots to be done to improve the reporting as per the reviews and [redacted]'s note. Again, how much do we want to delve in to the	Difficult one, but we will focus on the methods and the results for this	No change to manuscript

	battle for the data within this piece - they do go into that with the accompanying analysis. I'm not sure what the right answer is. As these will be open access and stand alone, I think we would want some of it within the paper and the discussion gives a nice summary of the struggle to get access to the reports.	systematic review	
	Report of the manuscript meeting of 6 February 2014 Updated By [redacted] - Research Committee on 06-Feb-2014 Present [redacted]		
3	* These papers are very important because they are going to be pored over and they may lead to legal action. We therefore feel they must be complete and live up to standards we expect for other systematic reviews and meta-analyses. Otherwise we worry that, having criticized industry for not providing detail you will be criticized for not providing it yourselves. We felt unable to appraise the statistical aspects of the paper because the data were not there -- all we had was the summary of results tables.	We will provide the statistical data and the forest plots they are substantial, and we did not provide these in the first instance due to concerns over space	Added in the forest plots and statistical data across the results section
4	* In some ways we thought this was the weaker paper of the two since the analysis of all of the data doesn't make a lot of difference. it does not really change our assessment of this drug. Nonetheless, we couldn't know that until the data were all evaluated, and you might wish to make that point in the discussion.	OK – interesting thought though that this drug is still recommended for stockpiling – there are also a number of new findings across the paper - before	Add to discussion
5	* We agree with the reviewer who suggests that the matter of harms needs more attention.	We will do this by adding in the assessment the figures and the analysis as per point 5	
	IMPORTANT		
	When you revise and return your manuscript, please take note of all the following points. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision.		
6	a. In your response to the reviewers and committee please provide, point by point, your replies to the comments made by the reviewers and the editors, and please explain how you have dealt with them in the paper. It may not be possible to respond in detail to all these points in the paper itself, so please do so in the box provided	We have outlined point by point responses	
7	b. If your article is accepted it will then be edited, proofed, and - after your approval - published on bmj.com with open access. This open access Online	Amended	

	First article will not be a pre-print. It will represent the full, citable, publication of that article. The citation will be year, volume, elocator (a unique identifier for that article): eg BMJ 2008;337:a145 — and this is what will appear immediately in Medline, PubMed, and other bibliographical indexes. We will give this citation in print and online, and you will need to use it when you cite your article.		
9	c. Please write an abridged version of the article for the print and iPad BMJ using the appropriate BMJ pico template for your study's design. Please be reassured that it doesn't take long to complete this. When your BMJ pico is ready please email it to papersadmin@bmjgroup.com.The templates for you to download are at http://resources.bmj.com/bmj/authors/bmj-pico	Amended	
	Please include the items below in the revised manuscript to comply with BMJ style:		
10	* the title of the article should include the study design eg "a retrospective analysis of hospital episode statistics"	Amended	Title incudes design – 'Zanamivir for preventing and treating influenza in healthy adults and children: systematic review of clinical study reports'
11	* ID of ethics committee approval and name of the ethics committee/IRB; or a statement that approval was not required (see ttp://resources.bmj.com/bmj/authors/editorial-policies/guidelines)	Amended	Added in on Page_ 10
12	* Please complete the following statement and insert it in your manuscript: Competing interests. Please complete the following statement and add it to your manuscript: "Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work [or describe if any]; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years [or describe if any]; no other relationships or activities that could appear to have influenced the submitted work [or describe if any]." (Please see http://bit.ly/T7uoG2)	Amended	Page 11
13	(see http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship) * copyright statement/ licence for publication (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)		Add in statement
14	* signed patient consent form(s), if the article gives enough personal	Not applicable	

	information about any patient(s): this sometimes occurs even in research papers - for example in a table giving demographic and clinical information about a small subgroup in a trial or observational study, or in quotes/tables in a qualitative study - http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/patient-confidentiality		
15	* for a clinical trial, the trial registration number and name of register – in the last line of the structured abstract	Not applicable	
16	* for any other registered study (eg a systematic review), the registration number and name of register – in the last line of the structured abstract	Not applicable	
17	*a data sharing statement declaring what further information and data your are willing to make available. Suggested wording: "Data sharing: technical appendix, statistical code, and dataset are available at this repository or website <url> OR from the corresponding author at <email address or url>". If there are no such further data available, please use this wording: "Data sharing: no additional data available"	We will be posting the CSR for open access – hopefully on the BMJ site	
	* please write the discussion section of your paper in a structured way, to minimise the risk of careful explanation giving way to polemic.		
	Please follow this structure:		
18	* statement of principal findings of the study	Amended	Page 8
19	* strengths and weaknesses of the study * strengths and weaknesses in relation to other studies, discussing important differences in results and what your study adds. Whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses (eg Cochrane reviews)	Amended	Page 9
20	* meaning of the study: possible explanations and implications for clinicians and policymakers and other researchers; how your study could promote better decisions	Amended	Page 10
21	* unanswered questions and future research	Amended	Page 10
22	* please note, too, that the article's introduction should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now	Currently two paragraphs	
23	* What this paper adds/what is already known box (as described at http://resources.bmj.com/bmj/authors/types-of-article/research)		Add in
	* funding statement (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements)		Page 10
24	* statement of the independence of researchers from funders (see	The views and opinions expressed	Page 10

	http://resources.bmj.com/bmj/authors/article-submission/article-requirements)	therein are those of the authors and do not necessarily reflect those of the Department of Health	
25	* for studies funded or sponsored by industry (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements):	Not applicable	
26	* a statement describing the role of the study sponsor(s), if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication	This project was funded by the NIHR Health Technology Assessment programme and will be published in full in the Health Technology Assessment journal series. Visit the HTA programme web site for more details www.hta.ac.uk/2352. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health	Page 10
	* assurance, in the cover letter, that a clinical trial funded by a pharmaceutical or other commercial company follows the guidelines on good publication practice (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements)	Not applicable	
27	* inclusion in the list of contributors the name(s) any professional medical writer(s), specifying in the formal funding statement for the article who paid the writer. Writers and authors must have access to relevant data while writing articles.	Not applicable	
28	* structured abstract (see http://resources.bmj.com/bmj/authors/types-of-article/research)	Amended	
29	* summary statistics to clarify your message We do want your piece to be easy to read, but also want it to be as scientifically accurate as possible. Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate: For a clinical trial: • Absolute event rates among experimental and control groups • RRR (relative risk reduction) • NNT or NNH (number needed to treat or harm) and its 95%	We have added the Absolute RD and the NNT or NNH where applicable We have also added these to the summary of findings tables	

	confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000) For a cohort study: • Absolute event rates over time (eg 10 years) among exposed and non-exposed groups • RRR (relative risk reduction) For a case control study: • OR (odds ratio) for strength of association between exposure and outcome For a study of a diagnostic test: • Sensitivity and specificity • PPV and NPV (positive and negative predictive values)		
30	As well as submitting your revised manuscript, we also require a copy of the manuscript with changes highlighted. Please upload this file with file designation 'Revised Manuscript Marked copy'.	Track changes document uploaded	
	REFEREES COMMENTS		
	Reviewer: 1 Recommendation:		
	Comments:		
31	Thank you for letting me read this innovative meta-analysis. It is well written and reported. I would have liked a bit more reflection on whether this enormous task actually changed anything for patients. I acknowledge that more data gives more assurance in estimates but it would be an even stronger case if the additional data changed the estimates.	Thank you Similar point to others that we outlined our error in trying to make the paper too succinct	Added in across the results see the track changes document
32	I am a bit concerned with the way harms are handled in this review. Even though it is general problem for meta-analysis this meta-analysis could potentially have solved it, given that the authors had access to clinical study reports, which usually consists of detailed information about harms. Three adverse events have been chosen without any motivation. Serious adverse events are not reported and neither are total number of adverse events. Based on the presented material it seems unsupported to claim that "Zanamivir tended to well tolerated". An expression that PRISMA considers "poor reporting practice".	We have added in the harms data and the relevant figures to the results section	Added in across the results see the track changes document
33	Your hypothesis that zanamivir may be no better than "relief medication" is interesting but since it was not a formal objective I believe that this reservation should be mentioned in the abstract. Perhaps it should also be mentioned that data was not pooled for this analysis. The entire analysis is	Given the reservation we have reported the pooled data in the results	Added in across the results see the track changes document

	a little bit odd. It is not a subgroup analysis but a comparison between all patients and then the subgroup who took relied medication. The outcome is different than the main outcome and it is not clear why only 8 studies were included (perhaps because only these studies reported on medication?). I think the entire analysis confuses more than it clarifies and I suggest that you remove it. A reference to the effectiveness of parcetamol or aspirin could be provided instead of the comparison.		
34	The relative reduction of prophylactic treatment is quite large but since the incidence of the disease is low the absolute reduction is small. If one could identify patients with a higher risk treatment might be warranted. However, on table 5 there are huge discrepancies between zanamivirs effectiveness in the symptomatic (RR 0.39) versus the asymptomatic (RR 0.97) and I suspect that this could be caused by lack or loss of blinding. A patient on active treatment might be less likely (or their physician) to classify their symptoms as influenza. Without knowing each individual study I would think that "asymptomatic" was a more objective measure of virus for instance in the nasopharynx and hence the RR of 0.97 might be the actual effectiveness of zanamivir.	I put this to the a whole team as it is a very important point We have considered it and given the inconsistency across results we have also added to the discussion that this inconsistency undermines the symptomatic result. The effect could therefore be, as the reviewers suggested, down to bias , the main	
35	Why are data on treatment reported for children and adult separately whilst data on prophylaxis are pooled?	For treatment it is straight forward to run adult and children trials, but in prophylaxis it is not straightforward. For example, post exposure prophylaxis trials in households will randomly allocate treatment to adults and children at the same time – thus leading to pooled prophylaxis results.	
36	I am curious as to whether the data supplied by GSK will be made publicly available and I am sure that the BMJs readers are as well.	The CSRs will be made available in full –having received them without a contract	
	Minor comments		
37	Some places it gets a bit confusing whether the results relates to children, adults or both. Perhaps subheading could solve this.	We have added in subheadings to the review to guide the readers We have done subheadings by complication as we perceive this is the optimal way to read the review – of	Amended SOF tables

		we do it by children adult we will end up with the converse problem - We have, however amended the SOF tables to make the children and adult sections clearer	
38	Asthma was a harms outcome in your Cochrane review why was it removed?	Check -	
39	The forrest plots are missing and I understand it is probably due to space restrictions. However it is problematic that I^2 and meta-analysis type (fixed or random) are not reported as this can have consequences for the interpretation of the results.	Amended	As per point 5 we have added in the forest plots and statistical data across the results section
40	Missing a 0 after 0.6 (0.60) in the tables and abstract.	Amended	Illness by 0.60 days (95% CI: 0.39 to 0.81 P<0.00001, I2 = 9%),
41	p4l17 I suggest a reference to support this statement.	Amended: This statement reference refers to the previous cited references in the sentence before: references 1 and 2	Added these references in a second place to the text to avoid confusion
42	p5 Perhaps you should explain the motivation for using these 3 extra criteria as they deviate from the usual Cochrane method.	Amended	Methods now reads as: Because of the sizeable quantity of available data
43	p5l22 Does that mean that you used fixed effect model for all other outcomes?	We used the fixed-effect method of Mantel and Haenszel as a sensitivity analysis to supplement our primary analyses using the random-effects method of DerSimonian and Laird. Random-effects meta-analysis is known to be overly conservative with sparse data. Hence we conducted sensitivity analysis using Peto's method on two occasions where we	methods now read as: To estimate treatment effects we first calculated the risk ratios (RRs) and used the average control event rate and the pooled RRs reported in the figures to calculate the risk differences (RD). For consistency we adopted this method for both the summary of finding tables and for the RDs reported in the text. For the analysis we chose to report

		had sparse data and borderline statistically significant results (prophylaxis with oseltamivir: renal body system on-treatment and psychiatric body system on-treatment). there were no endpoints for zanamivir that were marginally significant with sparse data.	the RRs as they are more consistent across the studies, and we have reported the heterogeneity for the pooled RR. We used Tau2 (inverse variance method) and I2 statistic to estimate between-study variance as measures of the level of statistical heterogeneity and the Chi2 test to test for heterogeneity and Hhigh estimates of heterogeneity were investigated where possible. We used the random-effects approach of DerSimonian and Laird based on mean differences for analysis of time to first alleviation of symptoms. For all other outcomes we used the random-effects approach for binary data of DerSimonian and Laird, where Tau2 was estimated using the inverse variance method.
44	p5l41 It seem reasonable to exclude NAIA3003 but you have not mentioned "usual care" as an exclusion criteria. If the review should be read alone I think it should be added.	We have substantially updated the methods to reduce the confusion	Added to methods the following: We included previously healthy people (children and adults), excluding those will illnesses such as malignancy or HIV infection. We included only trials on people exposed to naturally occurring influenza with or without symptoms. We targeted the intention-to-treat (ITT) and safety populations as prior reviews form our groups discovered compelling evidence that the intention-to-treat-influenza-infected (ITTI) were not balanced between treatment groups. Also these estimates will be more generalizable to practice where routine testing for influenza is not common and often not available. We included studies where zanamivir was administered by any route compared with

			placebo during the period in which medication was consumed (on treatment) and during the follow-up period of the study (off-treatment).
45	p5l44 You refer the Cochrane Risk of Bias tool but this also operates with "unclear risk of bias" which you have not used in your table. It seems problematic since the reason you have scored sequence generation low is probably that it has not been well described. It also seems inconsistent that you exclude open label studies even thou you have three trials where both blinding of participants and blinding of outcome assessment has been marked as "high risk of bias".	We have outlined our reasons for not using unclear in a separate paper, but to be brief our reasons are: 'The availability of full clinical study reports decreased the uncertainty and allowed definitive judgments to be made. "Unclear" risk of bias became a more certain "low" or "high" risk of bias, or even certainty of bias. Certainty or low levels of uncertainty were recorded against instances where our expectations of having all relevant and consistent information available for our reviews. When the information was not available, our judgments changed because we found gaps in the availability of information and inconsistent information. Whether the full study reports represent an exhaustive and coherent source of trial narrative and data remains unclear.'	
46	p5l53-58 Since this was not your prespecified outcome I believe that it belongs in the discussion.	It is relevant to the hypotheses on symptom alleviation,	
47	p7l20-50 The two first paragraphs seem to belong more to the background/introduction that in the discussion.	Amended:	We have moved the first paragraph to the introduction as suggested the second paragraph has been rewritten but now comes after the summary of the results
48	p8l36 Could you explain what "certificates of analysis" is? Why were they missing. Were they redacted/censored by GSK?	The certificates of analyses are documents which contain the information on both the intervention and the placebo (colour, appearance, etc). They were not available in the	Added

		CSRs – we do not know the reason	
49	p8l56 What kind of decision making are you referring to? Wasn't your analysis plan pre-specified in your protocol?	Risk of bias	Added in: to methods
50	p9l14 Antibody response results have not been presented in this paper so perhaps you should leave it out or specify that it is from the original Cochrane review.	We have added them in to a more comprehensive results section	Added to results sections: Data on hospitalisations for the zanamivir studies were not reported
51	p9l17 You have not shown that zanamivir lacks effect on hospitalisation. Previous you write that there is insufficient evidence which is not the same.	OK	We have amended the wording to keep it consistent across the review
52	p12 Flowchart. At the end of the flowchart it seems illogical that you have n=30 then exclude 2 but still write 30 underneath (30 zanamivir trials).	Amended	See Revised figure 1
53	It is not clear to me why what "unknown" as a exclusion criteria means.	Refers to trials for which there are identifiers but no CSRs or published data	Added in to Figure 1 "*trials for which there are identifiers but no clinical study reports or published data"
54	p16l15 What is phase of study? Phase I, II, III?	Phase 0: Pharmacodynamics & Pharmacokinetics Phase 1: Screening for safety Phase 2: Establishing the efficacy of the drug, usually against a placebo Phase 3: Final confirmation of safety and efficacy Phase 4: Postmarketing study	
55	p21 Table 1: Table heading: Is confirmed the same as completed?	Laboratory-confirmed influenza	Added as footnote to Table 1 *Confirmed cases indicate the number of participants with laboratory confirmation of influenza †Participants were aged 13 years and above (aged 16 and above, or aged 18 and above in some centres)
56	p27 If you swap the columns "Sinuisitis in children" with "Ottis media in	Amended	We have split table 3 into 2: adults and

	children" the columns for adult and children will be orderd in the same way.		children
	Additional Questions: Please enter your name: [redacted]		
	Reviewer: 2 Comments:		
57	I read the systematic review by Heneghan et al on zanamivir (Relenza) for the prevention and treatment of influenza with great interest. This is a very important topic especially since there is a similar review on oseltamivir (Tamiflu) – a drug that has recently attracted lots of attention (http://www.bmj.com/tamiflu) – submitted to BMJ (cited in page 7 of the manuscript). Undoubtedly, this is a well designed and rigorously conducted systematic review which uses the totality of the available evidence as documented in detailed by the authors. Thus it is expected that the effect of publication bias should be minimal. Given the recent progress in the UK and EU (Goldacre & Heneghan, BMJ. 2014;348:g213; O'Dowd, BMJ. 2014;348:g13), the approach used by the authors deserves lots of credit.	Thank you	
58	I understand that the authors have also conducted a concurrent review for Cochrane, whose reviews/reports are usually much more extensive than those published as journal articles. Although I had this in mind while reviewing the current manuscript, I tried to treat it (the manuscript) as a research article submitted to BMJ independent of the Cochrane Review. Having said the above, I am making the comments below, which mainly pertain to the depth of information provided in the BMJ submission. Given the authors' experience in systematic reviews, I have no doubt that the authors have already addressed most of these issues in detail in their Cochrane Review but I feel that these should also be addressed in the BMJ manuscript.	Similar points to the previous in that we have addressed this by providing more comprehensive methods and also results and figures	See point 44
	My comments/suggestions are:		
66	1. In the Methods section, the authors should specify their endpoints.	Have updated the methods section to add in the extra details as per point 44 - We were concerned over space in the first submission and realize some sections of the review were too brief	Added to methods Our primary outcome measures for treatment studies were symptom relief, hospitalisation and complications. For prophylaxis studies our primary outcomes were influenza (symptomatic and asymptomatic, always with laboratory confirmation and influenza-like

			illness (ILI), hospitalisation and complications, interruption of transmission (in its two components, reduction of viral spread from index cases and prevention of onset of influenza in contacts) and Harms.
67	2. Accordingly they should specify what effect metrics they used for which endpoints. In the current version, there seems to be some confusion for the reader since different metrics are used (RD, RR etc) which are not always clear what they represent. Also, I would suggest that they are consistent in how they express RD: e.g. RD=1.80% but later RD=-0.06 which is -6% I guess (p. 6)?	We have updated the results to reflect this issue	Added to methods For harms analsyis we were limited by the frequency of occurrence of some adverse events. We therefore meta-analysed (1) all serious adverse events; (2) all adverse events leading to study withdrawal; (3) all withdrawals; (4) all adverse events within a clinical study reports defined body system; as well as (5) a small group of common adverse events as defined in the FDA drug label for oseltamivir. There were too few events to meta-analyse deaths, serious adverse events by body system and any events that had overall incidence less than 0.5%. We did not meta-analyse outcomes with less than 10 events in total. Where appliocable we conducted analyses separately for on-treatment and off-treatment periods.
68	3. The authors should specify in the Methods section, how they evaluated heterogeneity (I2, Cochran's Q, tau2 etc). Also, I would suggest that they report the 95% CI of the I2 metric whenever applicable.	We have amended the methods as per point 44 and reported the methods but we also now added in the figures which gives the p value	
69	4. P. 5: "High estimates of heterogeneity were investigated when possible". Did the authors perform a meta-regression or some other analysis in order to investigate heterogeneity? How do they define "high estimates" (i.e. what value of I2 did they use as a threshold)?	Amended We did into use a cut-off as this is an arbitrary we investigated all possible heterogeneity as stated where the data permitted	Added in to methods We used Revman version 5.2 for the analyses and the forest plots. We used the random-effects approach of DerSimonian and Laird based on mean differences for analysis of time to first alleviation of symptoms. For all other outcomes we used the random-effects approach for binary data of DerSimonian and Laird, where Tau^2 was estimated using the

			inverse variance method. Additional analyses were reported as 'post-protocol'. We also planned to use the fixed-effect method of Mantel and Haenszel as a sensitivity analysis to supplement our primary analyses using the random-effects method of DerSimonian and Laird. Random-effects meta-analysis is known to be overly conservative with sparse data. Hence we planned to conduct sensitivity analysis using Peto's method where we had sparse data and borderline statistically significant results. However there were no endpoints that met this criteria for zanmaivir
70	5. Please specify what NNT stands for (number needed to treat) and what NNT means as a metric.	Have amended	Added in NNT and NNH throughout the review
71	6. The authors mention that they used random-effects models for the synthesis, while they used fixed-effects in sensitivity analyses. However, only one type of estimate is reported in the Results section which probably corresponds to random-effects. The authors should also provide the results of the fixed-effect analyses, which can be included in the appendix if not substantially different from random-effects (which seems to be the case since I2 is rather low for most meta-analyses).	We have updated the methods as per point 44 and 66 Added to the following	To estimate treatment effects we first calculated the risk ratios (RRs) and used the average control event rate and the pooled RRs reported in the figures to calculate the risk differences (RD). For consistency we adopted this method for both the summary of finding tables and for the RDs reported in the text. For the analysis we chose to report the RRs as they are more consistent across the studies, and we have reported the heterogeneity for the pooled RR. We used Tau2 (inverse variance method) and I2 statistic to estimate between-study variance as measures of the level of statistical heterogeneity and the Chi2 test to test for heterogeneity and Hhigh estimates of heterogeneity were investigated where possible.
72	7. P. 5: Trial NAIA3003 was excluded from the meta-analysis. Please	We have updated the methods as per	Added to methods

	clarify the inclusion criteria that a study had to meet in order to be considered in the meta-analysis. Also, please explain why usual care in the control arm seems be an exclusion criterion.	point 44 and point 66 and also added this point as well	We included previously healthy people (children and adults), excluding those will illnesses such as malignancy or HIV infection. We included only trials on people exposed to naturally occurring influenza with or without symptoms. We targeted the intention-to-treat (ITT), and safety populations, as prior reviews from our group discovered compelling evidence that the intention-to-treat-influenza-infected (ITTI) were not balanced between treatment groups. Also these estimates will be more generalizable to practice where routine testing for influenza is not common and often not available. We included studies where zanamivir was administered by any route compared with placebo.
73	8. The manuscript would be more comprehensive with the inclusion of forest plots (for the relevant quantitative syntheses) either in the main text or as web-only material.	We have done this, as per previous points we were concerned over space restrictions, and now see that this was an oversight	
74	9. I agree with the authors' discussion on the generalizability of the findings to low or middle income countries. Perhaps the authors could briefly comment on whether there are in general differences in baseline risks, prognosis, treatment implementation or administration etc. between high and middle/low income countries (Panagiotou et al, BMJ. 2013 Feb 12;346:f707. doi: 10.1136/bmj.f707).	Although this is an interesting point we have not added to our current discussion – with all the addition the paper is already getting rather long	
75	10. P. 8: More explanation is needed about the sentence: "Knowledge of new potential biases accumulated during the review process". It is rather unclear what the authors refer to with this statement.	We have added in the list of these new biases in the text	Added in: There were variations in the reporting quality of the included studies (Figure 2). Only one study showed adequate randomisation technique, while 25 (89%) demonstrated adequate allocation concealment. Adequate blinding of participants and personnel was reported in only two studies, while 24 (86%) demonstrated adequate blinding of outcome

			assessors. In addition, almost half of the trials had selected reporting and reported outcomes not specified in the protocol provided. A number of trials were under recruited and a number used different relief medication within the same trial across different centre and we also noted several other items that were not included in all full clinical study reports: • Certificates of analysis for the intervention/placebo preparations. • Patient enrolment dates explicitly reported (only trial inception and cessation dates are given; in zanamivir trials these are partially redacted). • Explicitly reported date of trial unblinding. We frequently noted the statement "the database was authorized on xxxx" to identify the unblinding date but an explicit date is important to report. In some cases, the date of unblinding was reported but the actual date within the month was redacted. • Authorship and accountability for the writing of the clinical study reports. • Statistical analysis plans in some cases. • Patient consent forms (missing from most zanamivir trials). • Patient information form (missing from most zanamivir trials). • List of randomisation codes (variably included).

			• Case report form templates in zanamivir trials do not allow for determining who completes the form (patient or clinician). • Core data sheet
76	11. Please confirm that numbers in parentheses after the estimates are 95% CI (this information is missing from some estimates).	Added 95% CI to all outputs	
77	12. The authors discuss the advantages of their effort to obtain all relevant data from GSK. They also mention in the Introduction that "these inclusions have been undermined by the presence of publication bias etc." (p. 4). I think that it would be useful if the authors could discuss the differences between the effect estimates in the current review and the previous reviews in Cochrane and BMJ (whenever estimates on the same endpoint are available). This would give the readers the chance to better appraise the value of investigators and/or clinicians having access to the totality of evidence obtained through a variety of sources.	Will add in section to results relating to the previous estimates	
78	13. Please expand on your discussion of the results (p. 5) of bias risk assessment according to Fig 2.	Have added into the bias assessment	
79	14. It would be useful if the authors could provide the number of studies that they identified through the materials provided to them by the manufacturer but had not been published in the literature – if they have this information available. Alternatively, it would be interesting to have an estimate (e.g. using a Venn's diagram) of the overlap between the different sources (shown in Appendix 1) that the authors used to identify the eligible trials. However, if understand correctly (p. 4 – Selection of studies), only trials from unabridged reports were included in the current analyses. The reason for this was that there were discrepancies between the published and unpublished reports, something that the authors should specify in more detail. Also, they should clearly specify that the sample of the potentially eligible trials was not affected by this selection (i.e. there were no published trials that they were not included in the clinical study reports finally included in the analyses).	This is a big question and really needs a another review – there are a number of additional questions that also go with this inclusion the additional reporting of harms compared to the published literature – Therefore we think this is really important; however, we want to focus in this review on the effects	No change
80	15. P. 9: In the "Implications for practice and research" the authors state that they do not believe that further RCTs are warranted. This is a very interesting statement as discussed previously (Ferreira et al, BMJ.	We could provide an optimal interval size calculation as suggested	

	2012;345:e5913). I think this could be supported by providing the 95% prediction intervals (Riley et al, BMJ. 2011;342:d549) of the effect estimates in a new study for the evaluated outcomes, especially since the authors assume random-effects models in their meta-analyses. This becomes even more valuable, since the authors estimates are expected to be unbiased or minimally biased given their search strategy.		
81	In summary, I would like to congratulate the authors for the intensive work they did in order to collect and analyze all available evidence on the topic. I look forward to a revised manuscript.	Thank you	
	Additional Questions:		
	Reviewer: 3		
	Comments:		
82	This is clearly an original contribution to the literature, and will be of great interest to readers of the BMJ. The issue of the degree to which existing systematic reviews in relation to the neuraminidase inhibitors are biased has generated considerable interest and has become a test case for access to clinical trials data to guide policy and clinical decision-making. The manuscript would therefore be of interest to a wide range of readers, including clinicians, teachers and policymakers. It will also be of interest to patients and is likely to generate considerable media coverage.	Thank you	
83	The most pressing problem with the manuscript relates to it being a subsidiary publication, intended to appear in concert with a publication in the Cochrane Library. It is unclear to what extent the manuscript submitted to BMJ is intended to stand alone as a readable and complete explanation of the methods used and the results obtained, or whether the emphasis is to be more extensively on the issue of missing data and the differences (if any) between this systematic review and those published previously. The following points highlight areas of concern about the completeness of reporting, which may makes interpretation of the BMJ publication difficult, without consulting the Cochrane publication.	Similar problem to previous points See points 1,3, 43,44,66,67,71 and 72	
84	1. In the PICO template, at line 17, the deletion of “or death” is presumed to be confirmed. Is this to be published as an annexure or a web version? Some statements in this document are difficult to interpret without concrete examples, in particular: “Knowledge of new potential biases accumulated during the review process”.	Add in th elist of new biases	• Certificates of analysis for the intervention/placebo preparations. • Patient enrolment dates explicitly reported (only trial inception and cessation dates are given; in zanamivir trials these are partially

			redacted). • Explicitly reported date of trial unblinding. We frequently noted the statement “the database was authorized on xxxx” to identify the unblinding date but an explicit date is important to report. In some cases, the date of unblinding was reported but the actual date within the month was redacted. • Authorship and accountability for the writing of the clinical study reports. • Statistical analysis plans in some cases. • Patient consent forms (missing from most zanamivir trials). • Patient information form (missing from most zanamivir trials). • List of randomisation codes (variably included). • Case report form templates in zanamivir trials do not allow for determining who completes the form (patient or clinician). • Core data sheet
86	2. On pg 4 of 29, from line 42, this sentence is difficult to interpret: “Due to discrepancies between published and unpublished reports of the same trials, we decided to include only those trials for which we had unabridged clinical study reports as well as and information on reports of trials which were considered 'pivotal' (i.e. first or second-line evidence to regulators in support of the registration application)”. Which option was exercised? Were any studies included on the basis of being ‘pivotal’, but for which no “unabridged clinical study reports” were available? If so, which were they?	Amended by clarifying the trials	Added “We included trials in previously healthy people (children and adults), excluding those with illnesses such as malignancy or HIV infection. We included only trials on people exposed to naturally occurring influenza with or without symptoms. We targeted the intention-to-treat (ITT), and safety populations, firstly because prior reviews from our group

			discovered compelling evidence that the intention-to-treat-influenza-infected (ITTI) were not balanced between treatment groups; secondly because estimates from ITT populations will be more generalisable to practice where routine testing for influenza is not common and often not available."
87	3. The utility of the two-stage extraction, appraisal and analysis process is unclear. How many studies were included in Stages 1 and 2? Are the published results only to reflect the results of Stage 2? The criteria for inclusion in Stage 2 are also somewhat vague – on pg 5 of 19, at line 7, the criterion for internal consistency is that "All parts (for example, denominators) of the same clinical study reports/unpublished reports were broadly consistent". What constitutes "broad" consistency?	30 trials were included in stage 1 and 28 in stage 2. Reasons for excluding 2 studies from stage 2 are reported in the results section and in Figure 1 (flow diagram)	Revised Figure 1
88	4. In relation to the external consistency criterion, Appendix 1 is intended to describe "a comprehensive strategy for dealing with data which we know are missing at the trial level, i.e. unpublished trials (see Appendix 1) and unreliable published records which are a very concentrated summary of clinical study reports". However, Appendix 1 provides a rather different rationale: "Although this review focuses on the primary data sources of manufacturers, to check that there were no published randomised controlled trials (RCTs) from non-manufacturer sources, we ran electronic searches in the following databases". It is unclear whether any of the studies provided by the manufacturer were excluded because of inconsistency between sources ("as reported in regulatory documents, other versions of the same clinical study reports/unpublished reports and other references").	Figure 1 provides a clear picture if the flow of trials	We have added to the result text the following 28 of zanamivir were included in Stage 1. Of the 28 trials in stage 1 we included 26 RCTs in stage 2 for the analysis. Two were excluded from the meta-analysis because one was only a synopsis (NAI30020) and one compared zanamivir to usual care and not placebo (NAIA3003). We finally included 26 zanamivir trials: 14 on treatment in adults and 10 trials in prophylaxis Our attempt at collecting sufficient information from regulatory files to reconstruct missing clinical study reports also failed because the information appeared insufficient for a reliable reconstruction.
89	5. Elements of the data synthesis process described in the manuscript do not appear to be included in the Results presented (but may well appear in the Cochrane version). There is little comment on identified heterogeneity or what are considered to be "high estimates", or of the outcomes of investigations. It is unclear what 'post-protocol' analyses were	We have added substantially to the methods and the review we have made it clear which methods we have used and why	

	implemented. There are also no indications of the results of sensitivity analyses ("pg 5 of 19, line 24: "We used the fixed-effect method for sensitivity analysis to supplement our primary analyses").		
90	6. As indicated above, it is unclear whether the 30 studies initially included constituted Stage 1 or 2. Of more importance, given the focus on access to missing data, it would be really important to show clearly whether the 30 studies included after the first set of exclusions were those provided by GSK or not. In Figure 1, 14 studies are listed as excluded for a reason shown as "unknown". What was unknown? This is almost half of the excluded studies and needs more detail and comment. Were any GSK-provided studies excluded in this step?	Unknown studies	Please see footnote in Figure 1
91	7. The risk of bias assessment is presented in detail in Figure 2, but receives scant mention in the Results. It is unclear whether the assessment was considered as material in relation to which studies to include in the meta-analysis. In particular, the assessment that only 1 study reported adequate randomisation technique needs commentary in the Discussion.	Add in to the results	There were variations in the reporting quality of the included studies (Figure 2). Only one study showed adequate randomisation technique, while 25 (89%) demonstrated adequate allocation concealment. Adequate blinding of participants and personnel was reported in only two studies, while 24 (86%) demonstrated adequate blinding of outcome assessors. In addition, almost half of the trials had selected reporting and reported outcomes not specified in the protocol provided. A number of trials were under recruited and a number used different relief medication within the same trial across different centre and we also noted several other items that were not included in all full clinical study reports:
92	8. Tables 1 and 2 are difficult to interpret without more detail. In particular, there is no legend explaining the meaning of "confirmed" (presumably referring to laboratory-confirmed influenza). In Table 1, in relation to study NAIB2007, the meaning of the age ranges cited is unclear.	Ok	We have added a footnote for confirmed; as well as some detail for NAIB 2007 IN Table 1
93	9. There are some errors/omissions in Table 3: is the relative effect measure in relation to pneumonia in adults an RR? Is the RR for otitis media in children intended to be 1.00? Should the final column be labelled as NNT	Amended	We have amended Table 3

	or NNH (95%CI), as is done in Table 5? In relation to bronchitis in children, is that an NNT?		
94	10. In Table 4, the last line of the final column is inconsistent with the style.	OK	Amended table 4
95	11. On pg 6 of 19, the paragraph starting at line 13 is unclear: "Of note, in clinical study report NAI 30012, the Protocol Amendment 11 (dated 24 Nov 2000) applied to all sites and clarified exclusion criteria to ensure subjects with severe persistent asthma were no longer recruited into the study". The intention of this paragraph is not clear, and requires more commentary. Should this have appeared in the Discussion?	Ok	Moved to discussion; added little commentary (Carl?)
96	12. A key element in this systematic review is the extent to which the request for access to company-held, and unpublished, data in the form of complete clinical study reports has added to the reliability of the estimates of effect. However, it is unclear, wither in the Results or the Discussion, which data were newly acquired, which were exactly as reported previously and included in previous systematic reviews, and which (if any) were previously included, but were changed as a result of access to the full clinical study reports. The statement (pg 7 of 29, line 31) that "Establishing clinical recommendations based on less than half of the evidence is clearly inadequate" is compelling. However, is that the case? How did the additional data change the outcomes of the systematic review (in terms of the effects estimates, their confidence intervals, the number of trial participants, etc)?	This is an important issue but a substantial piece of work in its own right. We have submitted another paper for publication (undergoing peer review) which examines how access for full CSRs changes risk of bias assessment. The results of the full CSRs for the 28 included studies are all new. We have removed the statement on page 7 of 29, line 31. We did not set out to compare data from unpublished studies with published data	We have added that the use of CSRs for our analyses is new. We have also included a section on agreement or disagreement with other studies
97	13. The most pressing question, which is probably impossible to answer, is whether health authorities (regulators or policymakers) would have taken a different line in relation to zanamivir at the time of the 2009 pandemic. It may be useful to recount some of the arguments used by such authorities in 2009 and subsequently, and commenting on whether this review would be likely to alter the conclusions reached.	This will be covered in the analysis and associated investigation pieces – we are focussing on the results for this piece	
	Job Title: Senior Lecturer [redacted]		
	Reviewer: 4		
	Comments:		
98	This is a very important and thoroughly performed review of the effects of zanamivir for treatment or prophylaxis of influenza outcomes and complications in randomised placebo-controlled trials. A major strength of	Thank you	

	the study is that it includes information from unpublished clinical study reports not available in previous reviews on the topic to minimise publication bias. The review is performed using state-of-the-art scientific methods and transparently reported. I have only a few comments, which mainly address the assessment of risk of bias in included trials.		
99	1. Although the authors make a great effort to obtain unpublished information from clinical study reports by the manufacturer to reduce potential publication bias, they set stringent criteria for reports to be included in the analyses (completeness of reporting, internal and external consistency). I wonder whether with these criteria they again introduce bias due to selective (incomplete) reporting. Can you please comment on this?	Amended	We have added to the results section and the discussion on this point It is also worth noting, that to date there has been no publically funded trial of zanamivir, which given we know manufactured funded trials overstate treatment effects, is somewhat puzzling, given the extensive use and stockpiling of this drugs.
100	2. Risk of bias items are reported for all included trials, but details about how these were assessed are lacking. Can you please report details of the definitions and assessments for all risk of bias items?	Added information	Added “To address the problem of reporting bias, we ignored published trial reports and directed our attention to clinical study reports and regulatory information. Our problems in reviewing the copious material at our disposal were how to identify and analyse important details in the midst of thousands of pages of information and how to construct a coherent appraisal of large and complex trial programmes. In addition since we gained unrestricted access to full clinical study reports (apart from personal de-identifying redactions) we took the view that all information needed to judge risk of bias should be present. Therefore when this information was not available, we judged the corresponding risk of bias element as at 'high' risk of bias”
101	3. The sentence about randomisation on page 5 line 44-45 does not make sense. How can 25 studies have adequate allocation concealment, when only one (of these?) showed an adequate generation of random	We used the Cochrane tool as risk of bias; randomisation and allocation concealment are separate domains in	See point 45

	sequence? If the random sequence is predictable, adequate concealment is not possible. Please clarify.	this tool	
102	4. How many of the studies were judged to have adequate, inadequate or unclear blinding of participants/personnel or outcome assessors? Currently, you only report the number of trials with adequate blinding.	We did not use “unclear”. We only used low or high.	See point 45
103	5. Please report also pooled results for time to alleviation of clinically significant symptoms (Table 4).	Added this in as per previous points	Added in figures
	Please enter your name: [redacted]		

Dear Editor

Manuscript ID BMJ.2014.017753.R1 entitled "Zanamivir for preventing and treating influenza in healthy adults and children: systematic review of clinical study reports" which you submitted to BMJ,

We have outlined a response to the first 3 points and also point by point responses in the associated table

We have uploaded three documents

1. A revised manuscript with track change
2. A revised manuscript with no track changes
3. A revised PICO document

In terms of these first three points

1. Can you somewhere indicate what the pre-specified outcomes were for each study included in this review? As a naive reader I found myself wondering this, and as someone who has previously done clinical trials, I think it is important because it bears on the quality of the data. Information about pre-specified outcomes typically is gathered and recorded more systematically than that for non-pre-specified outcomes.

Response:

The methods states:
A post hoc analysis was undertaken after we discovered seven trials provided data on time to first alleviation of symptoms with and without relief medication. Each patient in the studies may or may not have taken relief medication during the trial. Alleviation of symptoms may have occurred while the patient was taking relief medication and the "standard" comparison was made using this scenario. However, an additional analysis used a stricter definition where alleviation of symptoms could only be achieved without the use of relief medication. For example, a patient may have achieved alleviation using relief medication after 5 days but took 7 days to achieve alleviation without the use of relief medication. The comparison we reported is for all patients where we used the stricter definition for the zanamivir group (alleviation without relief medication) and the less strict definition for the placebo group (alleviation with relief medication).

And
Additional analyses were reported as 'post-protocol'.

And
For harms analysis we were limited by the frequency of occurrence of some adverse events. We therefore meta-analysed all serious adverse events;

However, I have difficulty understanding the comment : 'pre-specified outcomes typically is gathered and recorded more systematically than that for non-prespecified outcomes, In the context of using CSRs reviews, and we are also not aware of evidence underpinning this statement. -

2. Related to this, I did not see a clear description anywhere of why you have chosen the outcomes you did, and what other outcomes you might have included but did not. Please give a clearer explanation of your choices, particularly since, as [redacted] has mentioned, the protocol he looked at does not seem to specify the analyses based on relief medications. The outcome of pneumonia worries me in particular, because it was a rare event and I did not see anywhere a description of the definition of pneumonia used in any of the studies, nor an explanation of whether events were adjudicated or whether confirmation was sought. Perhaps this is in the papers, or implied somewhere, and I have simply missed it. If so, others might miss it too and it bears more emphasis. I am unsure as to this issue of why we chose what outcomes.

The reason for choosing these endpoints is consistent across all of our reviews. We chose them because they are all clinically relevant in terms of treating influenza. In terms of the relief medications we have made this clear in the methods, but again, it is a very important clinical endpoint: One that clinicians mention all the time and want an answer to. It is staggering that to date it has not been asked for. Of note, Mike Clarke, it in his BMJ editorial in 2009, called for a publically funded trial, still not done, but if it is to be done, it would now include a control arm on relief medications.

We have added to the methods the protocols for these.
Review Protocol: The review protocol was first published in 2011 (Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of unpublished data. Cochrane Database of Systematic Reviews 2011 , Issue 1 . Art. No.: CD008965. DOI: 10.1002/14651858) and subsequent amendments were published in 2012 and in the current review (see Feedback/Review Amendments 16 May 2013) [6]
In terms of the pneumonia issue the reviewer is correct to be worried ad there was no definition within the trials. That is why we report it as "pneumonia was a self-reported investigator-mediated unverified outcome"

We had long discussion about this as the endpoint was not a pre-specified end point in any of trials and was also not a specified secondary outcome of the trials. Unverified means it wasn't adjudicated, self-reported means no clinician confirmation, apart from the 2 trials with an x-ray definition. Some of these issues are discussed in detail in the analysis piece and the work has to be seen in the context of the related publication package.

Yet, it is important to note the relative reduction in pneumonia (often quoted as a 2/3^{rds} reduction) was used extensively for the licensing and stockpiling decision, without anybody, until this point in time, asking what exactly does this endpoint mean. Within the cochrane review we have provided the screenshots of the CRFs, because this is such an important point, so readers can judge for themselves. The implication is end points taken from published papers, without this level of scrutiny, are misleading. This is indeed worrying, but hardly ever addressed within the context of a review based on published papers alone.

3. Most research papers and systematic reviews/meta-analyses include in the discussion section a mention of their limitations and urge caution in interpreting any results that are uncertain. I missed a full discussion of the potential problems and limitations of this review

We have in our review a section on the strengths and weaknesses; the publication is now over 4700 words and could getting increasingly longer. We have additional information in the Cochrane review, which adds to this. We stand by the results that are presented throughout the review and have given a point-by-point response to the additional below. We have also submitted to BMJ open a full paper on issue relating to our risk of bias assessment that we have been working on: 'Risk of bias in industry-funded oseltamivir trials: comparison of journal publications and unpublished clinical study reports" manuscript ID is bmjopen-2014-005253. We look forward to your response.

Serial	Comment	Oseltamivir Notes	Zanamivir Notes	Action	Change to A159 needed
1.	As with the other paper, I would like you to address the additional matters of how you chose the outcomes considered in this study, problems with the definition of pneumonia as mentioned in the other decision letter	The text says: "….of clinical interest into primary and secondary by indication…." We carried out a meta-regression looking at the effects of different data capture methods of the "pneumonia" outcome			
2.	and include a full discussion of the potential limitations to this review along with cautions about overinterpretation of the findings.	Strengths and weaknesses of the review is a whole paragraph in Discussion			
3.	You will find the statistician's comments at the end of this letter. In addition to [redacted]'s remarks, however, we would like you to take [redacted]'s comments on the other paper fully into account in this paper as well, since they have many matters in common. I have copied them below -- please recall they were written with regard to the zanamavir paper so that some specific comments will not apply to	[redacted] made no actionable comments: "The statistical analysis of the data is appropriate and the presentation of the results has been restructured to enable a clearer interpretation of the data. Furthermore, a		Thank you	N

	this paper; you may discount those. Otherwise, please revise this one in line with his suggestions.	fuller discussion of the limitations of the review is presented. I can find no major statistical issues with this revised paper".			
4.	In peer reviewing this document I have taken the view that it is of utmost importance to ensure that the science and reporting here is squeaky clean. Hopefully this review is the first of a generation of reviews of evidence freely provided by industry, and it seems to be of utmost importance that it is undertaken to the highest standards, particularly given the criticism that industry has received from the Cochrane Collaboration, these authors, and others, over their reluctance to make trial data publicly accessible.	We agree and thank [redacted]		No action	N
5.	I have also been provided with access to the full Cochrane Review from [redacted], editor of the Cochrane Library. My review is quite long and detailed – I hope that I have identified all the points which will allow this review to be improved and made ready for publication. I apologise if I have misread any issues.			No action	N
6.	Most major points 1) There is no mention of a protocol for this systematic review. I have presumed that it is based on the Cochrane Protocol published in 2011 Issue 1. Including this information in the review is important (PRISMA item 5). There are several ways in which I have noted that the review differs from the protocol: a. Protocol includes comparisons with placebo or standard care, the review only included comparisons with placebo.	Placebo is the only comparator in the parent A159 Cochrane review. [redacted]'s cited protocol is not the one we followed to do the review. The protocol point is well made	Added to Zanamivir text as well	Added to bottom of methods text: **Review Protocol** The review protocol was first published in 2011 (Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, Thompson MJ. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children - a review of unpublished data. Cochrane Database of Systematic Reviews 2011 , Issue 1 . Art. No.: CD008965. DOI:	POSSIBLY

				10.1002/14651858) and subsequent amendments were published in 2012 and in the current review (see Feedback/Review Amendments 16 May 2013[15])	
7.	b. Plans to correspond with the trial's sponsor or report authors where further information is required were included in the protocol but are not mentioned in the review.			Text has been edited to: "**Search strategy** A variety of methods applied to different sources (publications, registries, correspondence with manufacturers, and review of regulatory documents) were used to identify and retrieve manufacturer- and non-manufacturer-funded clinical trials and their clinical study reports"	N
8.	c. There are no plans for modifying the quality assessment process in the protocol	We used an extended custom-built version of the Cochrane risk of bias tool to appraise clinical study reports. This extraction sheet was finalized prior to but inadvertently not mentioned in our protocol amendments of May 2013.	We used an extended custom-built version of the Cochrane risk of bias tool to appraise clinical study reports. This extraction sheet was finalized prior to but inadvertently not mentioned in our protocol amendments of May 2013.	We have added this to the methods	N
9.	d. Results will reported in both absolute and relative measures	Comment not understood we report both	They appear as both in the SOF tables we spilt these into to two to ensure they were all reported	No action	N
10.	e. There is no mention in the protocol of the analysis by use of relief medications. Some of these points are developed further below.	L339: "finally, data on the effects of rescue or relief medication (mainly paracetamol /acetaminophen) were incompletely reported"	See point 16 – this point refers to the oseltamivir review Zanamivir methods has Additional analyses were reported as 'post-protocol'.	The sentence has been deleted	CHECK A159

			See responses outlined below		
11.	2) The authors describe a two-step process for selecting reports for inclusion in the review. This is poorly described in the review (a much better description is given in the protocol). Whilst I believe that I understand why this process was adopted (to ensure that the evidence included in the review was internally consistent), the process as described seems subjective and a weak methodology. At first read it appears to be based on assessing whether there is agreement between every document ever written on a study on every detail. I am not sure whether this is the case. Inclusion of a clearer list of the requirements would be helpful. In this review I am not clear whether any trial was actually excluded on this basis. One seems to have been dropped because no CSR was available, and a further one because the comparison was not with placebo (which makes me question why it was included in Stage 1 as it fails to meet one of the inclusion criteria, albeit one which has changed between protocol and review).	In our previous re-write we added: "The main text has been edited as follows: "Because of the novelty and size of clinical study reports we subdivided the extraction, appraisal and analysis of the data into a two-stage exercise. In Stage 1 we assessed the reliability and completeness of the identified trial data. This was particularly important in the early stages of the review when we had received incomplete clinical study reports and were unsure of the importance of the missing parts….". The abstract has been edited to: "We included 23 in Stage 1 (reliability and completeness screen) and 20 in Stage 2 (formal analysis)"	This shows that there is a current a lack of understanding of what a clinical report is and is highlighted by the confusion later in point 22 when referring to manufacturer reports in the cited BMJ review. Many of these reports, in stage 1 come with trial id numbers ie. NAIA2005. It is therefore impossible for the contents of some of them to be determined until the full study report is received and analysed. Our flow diagram clearly describes the number of studies identified, inclusion, exclusion and progression from identification to stage 1 to stage 2 of the review and reflects what we did. Because of the absence of trial programmes for zanamivir listing all sponsored trials completed or underway, we had to rely on a variety of sources for the reconstruction of the trial programmes and retrieval of relevant clinical study reports. This complexity is reflected in the flowchart,	We do not think this comment is applicable to this review No change for Zanamivir review	NO

			illustrating the study selection process for this review.		
12.	3) In the quality assessment judgements that were unclear because information was not reported were coded as being at high risk of bias. The Cochrane Risk of Bias tool indicates that these should be reported as unclear and not high risk. This change in the way the tool is being used was not specified in the protocol. As a reader I am very interested in knowing the difference between issues which are unclear, and issues which are clearly wrong. The importance of differentiating between poor reporting and poor method has been recognised for a long time, and it is unfortunate that a view has been taken by the authors that it should not be made in this instance. Some unreported issues are unlikely to be sources of bias - for example, the method by which the random order of allocations appears not to be described in the clinical study reports, but it is almost certain to have been done appropriately (as these trials would have been subject to many FDA and MHRA inspections and the thought that they would be done with alternate or other flawed randomisation methods frankly is unbelievable). I am thus really interested to know whether the CSRs did not report this issue, or whether there was evidence that it was actually done wrong. The text suggests that most of the red blobs in the quality assessment arise because of poor reporting - but I cannot tell this from what is reported in	Comparison with publications is outside the aims of the review	Important point bit beyond the scope		NO

	the review. It is also a shame that the authors have not contacted the company (the sponsor) for clarification as promised in their protocol.				
13.	4) The reporting of data on children is misleading as it largely focuses on statistical significance without reporting estimated effect sizes with confidence limits. The abstract and discussion simply say that the effect was not significant. In the text the treatment effect estimate is given as well. As the point estimate is actually greater than that for adults, simple reporting on the basis of P>0.05 is giving a partial and misleading impression of the findings to a reader. There is no suggestion here that the effect in children is any less than that in adults, which is not what is implied by simply stating the effect in children is not significant. The emphasis on reporting differences simply as "significant or non-significant" without stating effect estimates recurs elsewhere in the review (e.g. interpretation of reduction in asymptomatic influenza) and should be checked through.	This seems to be an abstract problem			NO
14.	5) The results for binary outcomes are presented in the abstract as risk differences and numbers needed to treat. The primary statistic used for meta-analysis was the risk ratio which is not mentioned in the abstract. The risk differences and NNTs have been obtained by applying the risk ratio to the "average" placebo group event rate (not stated whether this is a mean or a median, or whether weighting was used in its calculation). The results section states the risk ratios as well as the RDs and NNTs, but	All the required data are reported in the Tables	Based on a mean of the control event rates	Added in mean to the paper Added to oseltamivir review The following text has been added: "Relative risks and risk differences were used to estimate treatment effects for binary data and mean differences for time to first alleviation of symptoms. To estimate treatment effects we first calculated the risk ratios (RRs) and used the average (mean) control event rate and the	CHECK A159

	does not state the prevalence figure at which RDs and NNTS are computed. The discussion states that influenza was only reduced by a small amount, quoting the NNT, and then questions whether the marketing authorisation was justified based on this.			pooled RRs reported in the figures to calculate the risk differences (RD). For consistency we adopted this method for both the summary of finding tables and for the RDs reported in the text. For the analysis we chose to report the RRs as they are more consistent across the studies, and we have reported the heterogeneity for the pooled RR"	
15.	Whilst I agree with the authors that it is most important to present absolute effects, the framing in this instance is unfortunate. Relative risks tend to make effects look large, whilst in a prevention scenario risk differences will always be small if few people in the study develop the disease, even if a treatment was 100% successful.		We report a 55% reduction in prophylaxis (line 394)	No change	NO
16.	Let's look closely at the results reported in Figure 7 and on page 7 lines 27-29. The text states "Zanamivir significantly reduced the risk of symptomatic influenza in individuals, RR = 0.39 (95% CI, 0.22 to 0.70, I2 = 45%); RD = 1.98%, 95% CI: 0.98 to 2.54, NNT = 51 (40 to 103) (Figure 7)". The RD and NNT figures are profiled in the abstract, and the NNT figure in the main paragraph in the discussion (page 9, lines 25-34).		See later points	The comment is not applicable to this review	NO
17.	The RR of 0.39 indicates that 61% (nearly two thirds) of influenza cases are prevented by using the drug. In the placebo groups in the 4 trials included here 86 influenza cases were observed in 2644, a rate of 3.26%. Applying the relative risk reduction to this figure would predict that 1.27% would develop influenza if they were taking the drug. The difference		They all appear with this level of detail in the summary of finding tables Page 9 lines 25-34) to make these points more clearly. See Serial 20	The comment is not applicable to oseltamivir review	NO

	between these figures is 1.99% the risk difference, and the inverse of this gives the NNT of 50. Thus the main reason why the risk difference is small is because few people in these studies developed influenza when taking placebo, not because the drug is useful. I believe that it would be most useful for all these figures (i.e. 61% of cases prevented, reducing event rates from 3.26% to 1.27%, giving an NNT of 50 – but all with confidence intervals) to be reported so that a reader can understand that zamamivir does prevent influenza, but in a prophylaxis situation few people develop influenza regardless of whether they take the zanamivir, and hence absolute benefits are low. Notably the difference in the household analysis mainly occurs because influenza is more common (78 cases from 410 is a rate of 19%, nearly six times as high) and not because the effect of the drug is greater. The authors might want to reflect on the wording of their conclusion statement (Page 9 lines 25-34) to make these points more clearly.		See for zanamivir serial 26 on the point of preventing influenza		
18.	6) An important analysis reported in the results is that of the impact of zanamivir when given with relief medications. This analysis was not mentioned in the protocol, and I am very confused as to how it has been undertaken. The nature of relief medications is not explained. The abstract mentions "the effect of zanamivir was attenuated by symptom relief medications" and gives figures which it implies are comparing "placebo with relief vs zanamivir without relief". There is no explanation in the methods section contains of		Added to methods This is a post hoc analysis undertaken after we discovered 7 trials provided data on time to first alleviation of symptoms with and without relief medication. Each patient in the studies may or may not have taken relief medication during the trial. Alleviation of symptoms may have occurred while the patient was taking	The comment is not applicable to the oseltamivir review The comment is not applicable to oseltamivir review. See also serial 8 (6). No analysis of the effect of relief medication was attempted because the data were inconsistently reported across the CSRs We have added to the zanamivir methods	Add to A159 to the methods

	these analyses or how the comparison was constructed. The results section mentions that there were seven trials available that allow the comparison reported in the abstract – the data is reported in Figure 5. It is not at all clear to me how these data were obtained and how the use of relief medication was determined. For example, taking the results for study 3008, the characteristics of included studies table (Table 1) states this study had 262 on zanamivir and 263 on placebo. Table 4 reports median values for all participants and participants who did not use relief medication (but does not state how many are in this latter group). Figure 5 gives the total sample sizes but implies that it is using the mean value from the zanamivir group who did not use relief compared and the mean value from the placebo group who did use relief. But everybody in the trial appears to be included in this analysis. There is no division into mutually exclusive groups of those who used relief medication and those that didn't. This can't be right.		relief medication and the "standard" comparison was made using this scenario. However an additional analysis used a stricter definition where alleviation of symptoms could only be achieved without the use of relief medication. For example a patient may have achieved alleviation using relief medication after 5 days but took 7 days to achieve alleviation without the use of relief medication. The comparison we report is of all patients where we used the stricter definition for the zanamivir group (alleviation without relief medication) and the less strict definition for the placebo group (alleviation with relief medication). We believe this is a valid comparison of zanamivir without relief medication and placebo with relief medication	A post hoc analysis was undertaken after we discovered 7 trials provided data on time to first alleviation of symptoms with and without relief medication. Each patient in the studies may or may not have taken relief medication during the trial. Alleviation of symptoms may have occurred while the patient was taking relief medication and the "standard" comparison was made using this scenario. However, an additional analysis used a stricter definition where alleviation of symptoms could only be achieved without the use of relief medication. For example, a patient may have achieved alleviation using relief medication after 5 days but took 7 days to achieve alleviation without the use of relief medication. The comparison we report is of all patients where we used the stricter definition for the zanamivir group (alleviation without relief medication) and the less strict definition for the placebo group (alleviation with relief medication).	
19.	In the discussion the authors state (Page 9 line 4) that "symptoms may be prolonged in the treatment arm when compared to the placebo group on relief medication". It is completely unclear to me where the data to support this statement comes from. The same argument continues in the Implications for practice and research (Page 11 lines 20-22).		Prolonged We have removed prolonged and changed to However, further analyses reveal the effect upon symptoms is synergised by the	The comment is not applicable to Oseltamivir Changed zanamivir paper discussion to "However, further analyses reveal the effect upon symptoms is synergised by the use of relief medication, revealing symptoms may	Check a159

			use of relief medication, revealing symptoms may be no better in the treatment arm when compared to the placebo group on relief medication. Ultimately this effect warrants testing in an open label trial of Nis compared to relief medications.	be no better in the treatment arm when compared to the placebo group on relief medication." Removed the word "prolonged"	
20.	And isn't this analysis an observational rather than subgroup comparison? And is the use of relief medication determined by the participant in response to how they are feeling? There are multiple ways in which such a comparison could be severely biased (including being driven in the zanamivir group by response to the drug) which need to be acknowledged. I am not convinced that the authors' conclusions for these analyses are justified (and certainly the analysis as presented is erroneous).		The use or not of relief medication, and its subsequent analysis, from a clinician perspective, is an important endpoint to understand. This is a point that we have discussed for some time now, and the trial is, as stated, subject to confounding. In an unconfounded trial the only difference would be the intervention under investigation. This is not the case. Because symptoms do differ between the groups, then the additional use of relief medication introduces an important bias in the analysis of the symptom relieving effect. And yes, zanamivir patients could have taken less relief medications. The only way round this is to standardize relief medications	The comment is not applicable to oseltamivir review No change to zanamivir review	NO

			(not done by any trial) or compare the interventions on and off relief medications (not done by any trial). This is such an obvious analysis that I cannot understand why no one has asked for it before. Given also that the symptom relieving effects are censored. This fact was noted by the FDA as potentially also misleading in assessing symptoms We therefore perceive this is an important analysis, and has been omitted by the trials to date, and has been difficult to delineate until now. These drugs do have symptom relieving effects, which are likely to be comparable to other symptom relieving drugs. Therefore this important analysis should stay in. Not least because it will change how these trials are done and reported in the future. The purpose of the trial is to refute the null hypothesis, our analysis aids the interpretation of whether these drugs should therefore be stockpiled based on their symptom relieving effects.		

21.	Less major points 7) The phrase "clinical study report" is not widely understood (I have tested the phrase on several colleagues who did not know what was meant by it). The authors should find an alternative phrase or give a better explanation.	Clinical study report is an official ICH term. In the last 5 years the term and content has received extensive coverage. The text in the Introduction is supported by reference 9 which is an open access exploratory review of CSRs of 14 different drugs. Because of space restrictions I am not sure we can do more.	It is standard terminology, it would therefore be wholly wrong to change this – The reason it is not well understood reflects the novelty of the review.	We have inserted further explanation in the introduction: "In the case of oseltamivir clinical study reports mean length is approximately 1305 pages (median around 900 pages)".	CHECK A159
22.	8) I would have liked to read a clearer description of the search strategy in the text. Also, the paragraph explaining the inclusion criteria (Page 2 line 56 to Page 3 line 18) would be easier to follow if it organised the criteria according to the PICO elements plus description of the study design, as in the Cochrane Protocol and Review. It is not clear what the implications from the comment about "pivotal trials" (Page 3 line 2) – did these not have clinical study reports?	The relevant text says: "There was a mix-up with follow-up cards in the "pivotal" trials WV15670, WV15671 and WV 15730 which does not allow drawing any conclusions on the durability of symptom relief[21]".. and "Information on a problem with follow up cards in three "pivotal" treatment trials was only discovered thanks to FDA SBA papers". The implications are well described	, in addition to information on reports of trials which were considered 'pivotal' (i.e. first or second-line evidence to regulators in support of the registration application).	The headings "types of studies" etc have been inserted in the text. Does not apply to zanamivir review	Check a159
23.	9) The report refers three times to Appendix 2, each time seemingly to a different Appendix 2, none of which are provided.	Appendix 2 (Searches for Regulatory information) was provided	Added in additional appendix for zanamivir review see appendix 2	The comment is not applicable to oseltamivir review	NO

				Change to zanamivir	
24.	10) The extended quality assessment list (bullet point list reported on page 6 lines 37-50) is also far from standard assessment, and it is not clear what the value of this information is in assessing risk of bias. This reads more like the list of findings of an MHRA CTU inspection visit looking at execution according to Good Clinical Practice guidelines and Standard Operating Procedures than assessing items which are known to link to bias. These items were not specified in the protocol.	This comment has to be inserted in the context of the time it took us to get the complete set of clinical study reports (4 years). In 2011 we asked Roche a series a clarification questions which were not answered. We have carried out changes to the risk of bias tool and our methods of using it but they require a self-contained paper to report these.	In terms of published papers we do have overwhelming understanding of the important biases. We do not have this for CSRs, therefore they should stay in In addition the last round of reviewers asked for them Reviewers comments from previous round: P. 8: More explanation is needed about the sentence: "Knowledge of new potential biases accumulated during the review process". It is rather unclear what the authors refer to with this statement.	Previous reviewers asked for more information	NO
25.	11) The quality assessment criteria reported in Figure 2 are not described. I am particularly baffled by the "other bias" category. Some explanation of the domains in the methods or as a footnote to the figure would be welcome.	The BMJ Higgins paper categorises other bias as any other bias identified by the researchers which does not fit in the other categories. [redacted] is thinking publications, we dealt with clinical study reports. For example the presence of dehydrocholic acid in the placebo and the different coloured placebo cap are examples of types of other bias	The other biases are detailed in the Cochrane review risk of bias tables There has to be some understanding that these are summary reviews and the Cochrane review are the more comprehensive review and by following the logic the CSRs are the definitive review Seems to me it is logical to say, I can't see the results in the summary review, but if I	For explanation purposes we have added the text underlined The placebo capsules in oseltamivir trials contained dehydrocholic acid and dibasic calcium phosphate dehydrate (we classified these as other potential biases). Zanamivir we have kept the bullet list of other potential biases	NO

		which will only be detected in Clinical study reports	have more questions then go to the Cochrane review, and if it is not clear then read the CSR This is exactly why we set out to do it this way, driving more readers to the comprehensive reviews through the summary		
26.	12) The conclusion in the abstract and text does not mention the effect of zanamivir on prevention at all. This seems an oversight.		This is the effect Zanamivir significantly reduced the risk of symptomatic influenza in individuals, RR = 0.39 (95% CI, 0.22 to 0.70, I^2 = 45%); RD = 1.98%, 95% CI: 0.98 to 2.54, NNT = 51 (40 to 103) (Figure 7) as well as in post-exposure prophylaxis, RR = 0.33 (95% CI, 0.18 to 0.58, I^2 = 40%); RD = 14.84%, 95% CI: 12.18 to 16.55, NNT = 7 (6 to 9) (Figure 8). However, the heterogeneity of this effect was moderate, and apart from one study (NAI30034) the sample sizes were small. In contrast, zanamivir did not significantly affect the risk of asymptomatic influenza in prophylaxis of individuals, RR = 0.97 (95% CI: 0.76 to 1.24, I^2 =0%) (Figure 89) or households, RR = 0.88 (95% CI:	The comment is not applicable to this review	CHECK A159

			0.65 to 1.20, I^2 = 0%) (Figure 10). The important point is to realize that zanamivir has no effect on asymptomatic influenza, therefore this group is free to carry on infecting populations at large. Combining this with a very small effect upon symptomatic influenza means it will not prevent influenza outbreaks, given the infectivity and transmission rates of the influenza virus. while asymptomatic individuals shed the virus, a recent systematic review concluded that more studies are required to examine the transmissibility of infleunza in this group of individuals (Public Health Rep. 2009 Mar-Apr;124(2):193-6.). It is therefore debatable whether or not asymptomatics spread influenza. In addition a recent paper discusses that antipyretics might increase transmissibility of infectious diseases (Proc. R. Soc. B 2014 281, 20132570, published 22 January 2014). So the best position is that we don't know		

			and in the absence of knowledge about this, it remains very important to know something about whether or not NIs reduce spread of infection and the trials don't tell us. In the corresponding analysis piece in the BMJ the following is reported 'Nor was any trial properly designed to investigate whether the drugs can interrupt transmission of influenza virus. When used for prophylaxis, the drugs reduce the risk of developing symptomatic influenza, but whether either drug cuts the risk of asymptomatic influenza or transmission of influenza virus has yet to be properly investigated.'		
27.	13) The authors reports a comparison of treatment effects by infection status (Page 7 lines 18-22 and Figure 6) reporting a test for difference in subgroups. This nature of this test is not reported in statistical methods section. In contrast to the analysis by use of relief medications, this analysis does divide the participants into mutually exclusive groups according to their infection status.		See above point on subgroups	Comments not understood	NO
28.	14) The number of figures reported is		The last round of peer review	No change	NO

	excessive (18), and could easily be reduced by combining several plots into single figures. For example, the impact on different definitions of influenza and pneumonia could be put on single slides, similarly the harms.		comments asked for the opposite, e.g., point 73 in zanamivir comments 'The manuscript would be more comprehensive with the inclusion of forest plots (for the relevant quantitative syntheses) either in the main text or as web-only material.'	We are happy if editors want to redraw the figures and consolidate them	
29.	15) I would have liked to have a clear comparison of the results of this review based on the hard-won CSRs with that of previous reviews which used only data in the public domain. The authors state that not using all available data introduces bias but they do not tell us how it has affected the findings of this review.		This is a big piece of work and I agree it is important. It would take us some time to do and what is important for now is the result of the information we have now are presented to allow a clinical judgment of its use and stockpiling.		NO
30.	16) Page 9 line 37 - the BMJ has previously published a systematic review based on all clinical study reports supplied by industry – http://www.bmj.com/content/325/7365/619.	I thought no one knew what a CSR was	It is not clear if these were CSRs – I very much doubt they were full clinical study reports In our experience manufacturer reports can mean anything We obtained from the **manufacturer reports** from all industry sponsored randomised controlled trials that were completed by 25 May 2000 and that compared celecoxib with placebo or other NSAID in people with osteoarthritis and rheumatoid arthritis.		NO

Re: Manuscript ID BMJ.2014.017753.R2 entitled "Zanamivir for preventing and treating influenza in healthy adults and children: systematic review of clinical study reports"

Many thanks for your patience as I collected and synthesized comments from our editorial team about how we would like this paper revised. We appreciate your patience with the process and with us.

1. We've already discussed our desire to have at least some indication -- to the extent possible given time constraints -- about the outcomes that were pre-specified for each study. It's clear that these are more difficult to identify than we realized when we made this request of you. You've indicated that there are discrepancies among pre-specified outcomes in the protocols, amended protocols and other study documents that are available to you. Because trial registration requirements were not in force at the time these studies were done, there is no way to be certain which outcomes were pre-specified. You suggested you would make an effort to identify these to the extent possible and include this information in the table that lists the characteristics of included trials.
Response: We have added additional tables, as web appendices. Including pre-specified outcomes, by trial ID, for each study, this also includes the primary and secondary outcomes. We have also added the protocol amendments, as these are important, given many of them impact on the pre-specified outcomes, including the definitions. Added to methods text: *The detailed list of pre-specified outcomes and protocol amendments in zanamivir treatment and prophylaxis trials are available in web appendix 3 and 4.*
2. In our initial suggestions for revision we indicated our worry about the protocol for this review. We've since had discussions about this and understand that various amendments were made to the original protocol, some of them as a result of inspecting the CSRs. You mentioned that until you saw the CSRs you didn't know what data were available regarding, for example, some of the harms outcomes. We understand the difficulties you were up against, but to forestall criticism we would like you to be as transparent as possible in these papers about exactly how the review differs from the original protocol, and when and why each decision to deviate from the protocol or add to it was taken. So, for example, it would be good to give a clear accounting of when and why you decided not to correspond with sponsors or authors.
Response:

The protocol and what we have done, and the changes are an important point.

To do this with transparency, we are producing a detailed summary of all of our review responses to over 35 peer reviewers, also some of the detailed protocol methods that have led to changes over the time period of these reviews

We have summarized for the zanamivir review, and added to the paper, the methods now reads as:

We have made a number of changes to the systematic review during the process of turning the protocol into the review. We have re-written the objectives twice, tightening up the text to bring it in line with our initial intentions and clarifying its meaning. The old objectives were: "To review clinical study reports (CSRs) identified from published and unpublished randomised controlled trials (RCTs) and relevant regulatory data on effectiveness and harms of NIs for influenza in all age groups" and "To review published and unpublished clinical study reports and other relevant regulatory data on effectiveness and harms of NIs for influenza in all age groups (and compare them with our published review)." We changed the emphasis of the objectives on unpublished study reports as we had decided from the start to concentrate on regulatory information. Similarly, comparison of published versus unpublished data is an important and worthwhile effort, but the original objective possibly misled readers as to its importance in our work. We had always conceptualised it as a low priority task we could carry out only if we had time following our review of unpublished data. We have also avoided using acronyms, which we thought cumbersome and confusing to the reader.
Our initial intention was to review clinical study reports and regulatory comments making up what we have subsequently called 'regulatory information'. The edits do not reflect a change in intent but our slowly evolving understanding of the problems we faced and our solutions to address these problems. While the review was underway we became aware of a number of other biases that we judged and recorded. The extraction sheet, for the risk of bias was finalized prior to, but inadvertently not mentioned, in our protocol amendments of May 2013. A post-hoc analysis was also undertaken by mode of pneumonia diagnosis: in two zanamivir adult treatment trials pneumonia was based on a stricter definition of X-ray confirmation and in nine trials pneumonia was based on a self-reported investigator-mediated unverified outcome.

A further post hoc analysis was undertaken after we discovered seven trials provided data on time to first alleviation of symptoms with and without relief medication. Each patient in the studies may or may not have taken relief medication during the trial. Alleviation of symptoms may have occurred while the patient was taking relief medication and the "standard" comparison was made using this scenario. However, an additional analysis used a stricter definition where alleviation of symptoms could only be achieved without the use of relief medication. For example, a patient may have achieved alleviation using relief medication after 5 days but took 7 days to achieve alleviation without the use of relief medication. The comparison we reported is for all

patients where we used the stricter definition for the zanamivir group (alleviation without relief medication) and the less strict definition for the placebo group (alleviation with relief medication).

In terms of our correspondence we do not feel we have little to worry about:
Many of our requests are detailed on the BMJ open data campaign
http://www.bmj.com/tamiflu/roche

in terms of Relenza there is a similar page
http://www.bmj.com/open-data/relenza

As an example this includes accounts asking for SAEs, in addition to a number of other questions see email dated 14th May 2013 to [redacted] [GSK contact]

We have therefore added the correspondence details in the paper methods:

Inhibitors in children [7] and healthy adults [8] and then updated the searches again on the 22 July 2013. Our detailed correspondence with GSK, to access data, is available online at the BMJ open data campaign http://www.bmj.com/open-data/relenza

3. We remain very uneasy about the way in which you have assessed the quality of the CSRs and are not as persuaded as you seem to be that the adaptation of the Cochrane ROB tool was needed. As you know, a main point of difference is the decision you took to assign a high risk of bias in situations where allocation concealment was not described in the CSRs.

Several of us who have experience with industry-conducted trials have the experience that in industry-conducted trials these things are done very carefully. We think this is a matter of reporting quality and do not agree the risk of bias is high here. We can, however, accept that you disagree with this, but would like you to revise the paper to discuss this in terms of reporting quality rather than bias. Your discussion should acknowledge that these are separate issues, and reinterpret things in terms of reporting quality.

Another problem in your adaptation of this instrument is that we surmise it was developed with a knowledge of what was in the CSRs, whereas most such tools are developed or adapted for more general reasons. A critic might say that this process of adaptation is unfair to the companies and might have been set up for industry to fail.

Response:

To address this point we have added to the discussion:

In addition, we used the Cochrane seven-domain 'Risk of bias' instrument to assess bias. The availability of partial or complete clinical study reports decreased the uncertainty and allowed us to make definitive judgements. Previous unclear risk of bias therefore, became certainty of presence or absence of bias. However, there is still some uncertainty as to whether the complete study reports represent an exhaustive and coherent source of trial narrative and data, which would undermine this judgement.

In terms of industry trials are done "carefully", we would disagree with this point., particularly given the recent BMJ article on broken trials, which comes to an opposite conclusion:

See FDA official: "clinical trial system is broken" http://www.bmj.com/content/347/bmj.f6980
"The clinical trial system is broken and it's getting worse, according to longstanding Food and Drug Administration investigator, Thomas Marciniak."

"The latest is a trial by UK drug company AstraZeneca of the antiplatelet drug ticagrelor (marketed in the United States as Brilinta and in the European Union as Brilique).2
"Drug companies have turned into marketing machines. They've kind of lost sight of the fact that they're actually doing something which involves your health," Marciniak says. "You've got to take away the key components of the trials from drug companies."

4. The process for selecting reports to include in the review is not described in enough detail that it could be replicated. Some of the steps seem very subjective, particularly the methods for identifying consistency. Did you mean by this that all numbers had to match among the various documents, or that every subject had to have an outcome, or what? Please be much more detailed and try to describe the methods so that others could replicate this important part of the process.

This is important because you have taken the view that anything not included in the CSRs must have been left out or done improperly – uncharitable readers might take the same view of your own methods in this study.

Response:

We have added the following to the methods :

"2) Internal consistency. All parts (for example, denominators) of the same clinical study reports/unpublished

reports were consistent. Access to full clinical study reports allowed us to follow consistency across chapters and appendices, creating a need for far more interaction with the text. The parts of a clinical study report we checked for consistency included the core report, the pre study documents, study methodology, individual subject listings of demographic and efficacy data, and individual listings of safety data, as well as the statistical analysis plan and the serious adverse events."

5. We continue to think that the paper overstates things with regard to children. One of the editors involved made the remark that based on the findings here, in a pandemic situation the evidence seems strong enough that they would give the drug to their child. So we reiterate that we think you must acknowledge that the point estimate here is actually greater than that for adults. Simply interpreting this finding as significant or not does not acknowledge that the CI is compatible with benefit.

Response:

We disagree with the statement 'in a pandemic situation the evidence seems strong enough that they would give the drug to their child.' I think this statement needs clarifying, as it is not how we heard it, and it is very hard to understand given the lack of clinical impact reported in the results on children.

In terms of the effects of Zanamivir in children:

The number needed to benefit are very high and also not significant in treatment trials

Children						
Time to first alleviation of symptoms [days]		**1.08 lower** (2.32 lower to 0.15 higher)	1.08 days (-2.32 to 0.15)	723 (2)	N/A	N/A
Complications: Pneumonia	**12 per 1000**	**6/1000** (1 to 28)	**RR 0.53 (0.12 to 2.38)**	737 (2)	0.56% (-1.64 to 1.04)	**NNTB 178** (NNTB 96 to ∞ to NNTH 62)
Complications: Bronchitis	**15 per 1000**	**13 per 1000** (4 to 41)	**RR 0.86** (0.26 to 2.80)	737 (2)	0.21% (-2.67 to 1.10)	**NNTB 482** (NNTB 92 to ∞ to NNTH 38)

Complications: Sinusitis	**15 per 1000**	**13 per 1000** (2 to 96)	**RR 0.87** (0.12 to 6.45)	737 (2)	0.19% (-8.09 to 1.31)	**NNTB 519** (NNTB 77 to ∞ to NNTH 13)
Complications: Otitis media	**70 per 1000**	**70 per 1000** (42 to 122)	**RR 1** (0.59 to 1.72)	737 (2)	0% (-5.13 to 2.92)	**NNTB Infinity** (NNTB 35 to ∞ to NNTH 20)

Therefore we fail to see how the effects warrant use in a pandemic.

In Revman a test for subgroup difference between adults and children shows no evidence of a difference (Chi-square = 0.58, 1 df, P=0.45) with the overall effect being 0.66 days reduction in time to first alleviation of symptoms (95% CI 0.44 to 0.87) with I2 = 20%)

The point estimate does not warrant over interpretation when it is not significant. In terms of clarifying whether a point estimate is different to another based on just observation alone (ie., that it seems bigger), this is an incorrect methodological approach. The approach is, as above, to test for the interaction.

For a similar approach to using tests for interactions and interpreting subgroups (and including post hoc analysis)

See: Self-monitoring of oral anticoagulation: a systematic review and meta-analysis.
Heneghan C, Alonso-Coello P, Garcia-Alamino JM, Perera R, Meats E, Glasziou P.
Lancet. 2006 Feb 4;367(9508):404-11. Review.

We have therefore added the above results to the paper (page 8):

A test for subgroup difference between adults and children shows no evidence of a difference (Chi-square = 0.58, p=0.45) with the overall effect being 0.66 days reduction in time to first alleviation of symptoms (95% CI 0.44 to 0.87 days, I^2 = 20%)

However, in all of our reviews that we have done to date we have not been asked, nor been asked to overestimate, a non-significant point estimate.

6. This revised version of the paper seems to have added a statement that the drug does not work in asymptomatic persons, which we did not see in the original version of the paper. We aren't sure where that comes from.
Response: The paper states: Zanamivir significantly reduced the risk of symptomatic influenza in prophylaxis of individuals, RR = 0.39 (95% CI, 0.22 to 0.70, I^2 = 45%); RD = 1.98%, 95% CI: 0.98 to 2.54, NNT = 51 (40 to 103) (Figure 7) as well as in post-exposure prophylaxis of households, RR = 0.33 (95% CI, 0.18 to 0.58, I^2 = 40%); RD = 14.84%, 95% CI: 12.18 to 16.55, NNT = 7 (6 to 9) (Figure 8). However, the heterogeneity of this effect was moderate, and apart from one study (NAI30034) the sample sizes were small. In contrast, zanamivir did not significantly affect the risk of asymptomatic influenza in prophylaxis of individuals, RR = 0.97 (95% CI: 0.76 to 1.24, I^2 =0%) (Figure 9), nor in post-exposure in households, RR = 0.88 (95% CI: 0.65 to 1.20, I^2 = 0%) (Figure 10). The first version of the paper was amended at the request of a number of reviewers , Reviewer 1: "The conclusion in the abstract and text does not mention the effect of zanamivir on prevention at all. This seems an oversight." Reviewer 2: "The relative reduction of prophylactic treatment is quite large but since the incidence of the disease is low the absolute reduction is small. If one could identify patients with a higher risk treatment might be warranted. However, on table 5 there are huge discrepancies between zanamivirs effectiveness in the symptomatic (RR 0.39) versus the asymptomatic (RR 0.97) and I suspect that this could be caused by lack or loss of blinding. A patient on active treatment might be less likely (or their physician) to classify their symptoms as influenza. Without knowing each individual study I would think that "asymptomatic" was a more objective measure of virus for instance in the nasopharynx and hence the RR of 0.97 might be the actual effectiveness of zanamivir." We therefore have responded appropriately to review comments. We would also like to point out previous peer review comments that highlight: "The review is performed using state-of-the-art scientific methods and transparently reported."

Therefore our additions of the data are in line with responding appropriately to peer review.
7. We also think that this revised version of the paper does not give adequate emphasis to the finding that the drug seems to be effective for prophylaxis. Here the emphasis on the NNT is somewhat misleading. It is high because influenza is not common in this group, but despite this the drug seems to produce roughly 61% reduction in symptomatic cases. This seems to us a major effect, comparable to vaccines. We had suggested a sentence that could be included that describes this, and would like you to include it in the paper: "Zanamavir significantly reduced the risk of symptomatic influenza in individuals, RR=0.39 (95% CIO, 0.22 to 0.70, 12=45%); RD=1.98%, 95% CI: 0.98 to 2.54, NNT=51(40 to 103) (Figure 7)." The abstract should also also mention that this reduces event rates from 3.26% to 1.27%. This will help readers understand that the absolute benefits are low.
Response: We are happy to change to the RDs in the abstract, but will also add in the RDs for asymptomatic individuals The abstract now reads as: *Zanamivir tended to be well tolerated. In zanamivir prophylaxis studies, symptomatic influenza in individuals was reduced by a small amount, RD = 1.98% (95% CI: 0.98 to 2.54); reducing event rates from 3.26% to 1.27%, which means 51 people need to be treated to prevent one laboratory confirmed case (95 % CI, 40 to 103). In contrast, the prophylaxis effect on asymptomatic influenza cases was not significant in individuals, RD = 0.14% (95% CI, -1.10 to 1.10), nor in households, RD = 1.32% (95 %CI, -2.20 to 3.84). In households treated prophylactically there was an effect on symptomatic influenza, RD = 14.84% (95% CI: 12.18 to 16.55%), but this was based on only two small studies including 824 participants.* Also We are unclear what is meant by the drug is effective for prophylaxis and where this come from. There is big difference between stating an intervention is statically significant as opposed to it is clinically significant On this note, I refer you to the previous review comments: "The relative reduction of prophylactic treatment is quite large but since the incidence of the disease is low the absolute reduction is small. If one could identify patients with a higher risk treatment might be warranted. However, on table 5 there are huge discrepancies between zanamivirs effectiveness in the symptomatic (RR 0.39) versus the asymptomatic (RR 0.97) and I suspect that this could be caused by lack or loss of blinding. A patient on active treatment might be less likely (or their physician) to classify their symptoms as influenza.

Without knowing each individual study I would think that "asymptomatic" was a more objective measure of virus for instance in the nasopharynx and hence the RR of 0.97 might be the actual effectiveness of zanamivir."

In our discussion we state:
In zanamivir prophylaxis studies, symptomatic influenza was only reduced by a small amount: 54 people need to be treated to prevent one person from having symptoms of influenza infection; importantly reductions in asymptomatic influenza cases were non-significant. While asymptomatic individuals shed virus, a recent systematic review concluded more studies are required to examine the transmissibility of influenza in this group of individuals (11). Whilst it might be debatable whether or not asymptomatic spread influenza, the current results do not provide evidence of an effect upon asymptomatic and upon reducing the risk transmission.

The emphasis on the finding that the drug seems to be effective for prophylaxis cannot be justified based on the results and from a clinical perspective, particularly if you take the clinical significance into account. Clinical significance, from a population perspective, would include the size of the effect, how many people the drug prevented disease in, and for how long, given the length of exposure a population has during an influenza outbreak. Based on these parameters, the drug, although statically significant, is clinically irrelevant. Given also the lack of effect on asymptomatic and the lack of evidence on transmission we stand by the presentation of our results and the lack of effect.

8. The editors have had extensive discussion about the analysis by use of relief medications. We think this is hopelessly confounded with treatment assignment and do not feel it should be in this paper. We ask that this be removed. It might be something that you could examine in a separate paper. Instead of this analysis, we ask that you simply present the proportions in each group that took relief medications in the two arms of the trial so that readers can judge things for themselves.

Response:

I discussed this in person with [redacted] reviewer.

Sending this additional explanation
No, it is not a subgroup analysis. In the primary analysis all patients could be taking relief medication at the time their symptoms were relieved. In a "secondary" analysis alleviation could only occur if the patient was not taking any relief medication. Here is a hypothetical patient:

Time (days)	1	2	3	4	5

Relief (Y/N)	Y	Y	Y	Y	N
Alleviation (Y/N) N	N	N	Y	Y	

In the primary analysis alleviation occurred at 4 days but in the secondary analysis it occurred at 5 days. So for our post-hoc analysis we compared the zanamivir group using the secondary analysis definition (zanamivir without relief) with the placebo group with the primary analysis definition (placebo with relief).

It is an important endpoint, as it is clinically relevant to the whole mechanism. It will be in the Cochrane review and it would be wrong for us to remove it from the BMJ review, as it would mean we are being asked to selectively report an endpoint.

There seems to me no clear statistical reason as to why it should be removed.

In terms of our statistical review comments:
[redacted] made no actionable comments: "The statistical analysis of the data is appropriate and the presentation of the results has been restructured to enable a clearer interpretation of the data. Furthermore, a fuller discussion of the limitations of the review is presented. I can find no major statistical issues with this revised paper".

However, we have added to the discussion, in the limitations section, the following:

Finally, our analysis of relief medications only came about because a number of trials, but not all, reported data on alleviation of symptom whilst not taking relief medications. This analysis was not pre-specified in our original protocol, and therefore to confirm the symptom alleviating effect an open label trial of zanamivir versus standardized relief medications could be undertaken.

9. Perhaps we missed it but we could not find the appendix 2 for this paper that you mention. Might it be in the supplemental files for the other paper?

Response:

We will reload it

Of note there are now six appendices

10. In response to our previous comments about the “other bias” category, we would like much more description of some of the things that were considered here. You mention, for example, the presence of dehydrocholic acid in the placebo – these things should be put in the paper.
Response: We have put together a list of all the other biases as an appendices: See appendix 6 Added to results: *“A further explanation of these other biases can be found in Appendix 6.”* Use of dehydrocholic acid (a stimulant laxative) as placebo relates to oseltamivir Zanamivir is effected by lactose See: The placebo for zanamivir trials contained lactose powder, which can potentially cause bronchospasm, but certificates of analysis for the intervention/placebo preparations were not available except for one trial
11. In this revision you indicated that you did not understand our previous request that you should report the nature of the test used for the comparison of treatment effects by infection status. (Your ponit 27 in your response to the original decision letter.) What we would like is a statement in the methods section of how you did this. Was it, for example, meta-regression?
Response: Have added to methods: *We used the chi-square test for subgroup differences provided in Revman*
12. We also think you could do more to put the results of your review of CSRs in the context of previous reviews in the public domain. This would help substantiate the point you make that not using all available data has compromised previous reviews.
We have further added to the discussion the following: *Our review, by better clarifying definitions of endpoints, disagree with previous reviews [12] that reported the evidence consistently supports zanamivir as being clinically effective. None of the previous reviews were able to clarify how diagnostic endpoints were verified (i.e. pneumonia), and delineated self-reported outcomes from those that are verified by objective measures. [1, 2, 8, 12, 13] To our knowledge, no trial or systematic review has reported on the use of relief medications and analysed this important confounder. We have provided the*

most comprehensive and up to date analysis of harms and we also can confirm that data on hospitalisations for the zanamivir studies were not reported. Further to this we have reported a number of biases that are not seen in journal publications, but are relevant to the assessment of clinical study reports. In addition, our results do not provide evidence of an effect upon asymptomatic influenza and upon reducing the risk transmission.

I hope you will interpret these requests in the spirit in which they are intended, which is that we want the best possible paper to come out of this process!