Author's response to reviews

Title: Comparing influenza vaccine efficacy against matched and mismatched strains: A systematic review and meta-analysis

Authors:

Andrea C Tricco (TriccoA@smh.ca)
Ayman Chit (aymanchit@gmail.com)
Charlene Soobiah (SoobiahC@smh.ca)
David Hallett (hawk.hallett@utoronto.ca)
Genevieve Meier (genevieve.c.meier@gsk.com)
Maggie H Chen (maggiehchen@gmail.com)
Mariam Tashkandi (mariam.tashkandi@gmail.com)
Chris T Bauch (cbauch@uoguelp.ca)
Mark Loeb (loebm@mcmaster.ca)

Version: 3 Date: 23 March 2013

Author's response to reviews: see over
Dr. Claire Barnard, PhD  
Senior Editor  
*BMC Medicine*

March 23, 2013

Re: MS 2700060098655244: Effect of influenza vaccination against mismatched strains: A systematic review and meta-analysis

Dear Dr. Barnard,

Thank you for requesting a revised version of our manuscript. We thank the peer reviewers for their thoughtful comments, which we feel have strengthened our paper.

We have carefully considered each of the recommendations for improvement and have made revisions to the manuscript. Enclosed you will find our responses to each of the reviewer comments. Please note that the line numbers correspond to the document with tracked changes. As suggested in your email, a revised manuscript with tracked changes has been uploaded to your online system, as well as a cleaned version.

We look forward to receiving a final decision from your journal.

Best regards,

Mark Loeb, MD, MSc, FRCPC  
Faculty of Health Sciences, McMaster University,  
Michael G. DeGroote Centre for Learning,  
Rm. 3203, 1200 Main Street West,  
Hamilton, Ontario, L8N 3Z5, CANADA  
Telephone: 905.525.9140 x 26066, fax: 905.389.5822, email: loebm@mcmaster.ca
REVIEWER 1: Heath Kelly

1. Neither the abstract nor the title of the manuscript indicates to the prospective reader that the meta-analysis is of both matched and mismatched strains. The conclusion in the abstract that efficacy is higher for matched strains cannot be drawn from the data provided in the abstract. May I suggest a revision of the title along the lines of “A comparison of influenza vaccine efficacy against matched and mismatched strains”? Please add findings on matched strains to the abstract.

We agree with the reviewer and have changed the title to “Comparing influenza vaccine efficacy against matched and mismatched strains: A systematic review and meta-analysis”. We have also provided estimates of efficacy for matched strains in the abstract, as suggested.

2. Because the review is of RCTs, the authors generally refer to ‘efficacy’. Occasionally ‘effectiveness’ appears. Is the latter term intended?

We have changed the term “effectiveness” to “efficacy” on line 68. We believe that the word “effectiveness” is more appropriate on line 72, as this is referring to “real-world” experience versus RCTs. Effectiveness does not appear anywhere else in the text.

3. Estimates of the number of influenza-related deaths and hospitalisations in the US are subject to some debate. The rationale for this study would stand without the first sentence of the Introduction.

We have revised the first sentence as follows (line 56): “Influenza is a major public health threat.”

4. Please define a ‘quasi-RCT’.

This has been added to lines 93-94: “…quasi-RCTs (i.e., use of non-random methods to allocate patients to the treatment and control groups, such as consecutive enrolment or the last digit of a health card number)…”

5. Some further clarification of the definition of mismatch would be helpful. As the authors note, this has not been satisfactorily attempted in previous meta-analyses.
   a. In the methods it might be easier to describe B mismatches more simply as lineage mismatches or drifts within lineages.

This section has been revised in our methods to try and clarify the concept of mismatch. Beginning on line 138: “Influenza A strains from infected trial participants were matched with the strain in the vaccine if they belonged to the same A subtype (i.e., H1N1 or H3N2) and were antigenically similar in the HI assay (i.e., if they showed sufficient cross-reaction in a HI chessboard table using ferret antisera; for example, with a HI typing quotient <4-fold titre). Influenza A viral strains were considered mismatched by antigenic drift if they were antigenically distinct from influenza A strains contained in the vaccine as per HI typing (e.g., HI titre quotient ≥4-fold) or the characterization did not belong to a similar influenza A subtype.
contained in the vaccine (e.g., H1N1 strains circulating but only H3N2 strains contained in the vaccine for bivalent vaccines with one H subtype).

For influenza B, the epidemiological situation is complex. In recent years, there have been two co-existing phylogenetic influenza B lineages; B/Victoria and B/Yamagata [20, 21]. Influenza B strains from infected trial participants were considered matched if the strain belonged to the same lineage and were antigenically similar to the vaccine strain as per HI typing (e.g., HI typing quotient <4-fold titre). For influenza B mismatches, two different forms were considered. Mismatch by antigenic drift refers to strains of the same lineage that were antigenically distinct from influenza B strains contained in the vaccine as per HI typing (e.g., HI quotient ≥4-fold titre), whereas mismatch by lineage refers to influenza B strains of different lineages. Whenever the influenza B lineage was not presented in the trial report, categorization was based on the influenza phylogenetic tree and verified by influenza experts on the team.”

b. *I don’t think it is appropriate to call an H3 vaccine strain mismatched to an H1 circulating strain – if this is what is intended on p9. On re-reading perhaps this in intended only for bivalent vaccines with one H subtype. This would not apply to TIV. Can the authors please clarify?*

This has been clarified on lines 147-148: “...(e.g., H1N1 strains circulating but only H3N2 strains contained in the vaccine for bivalent vaccines with one H subtype).”

c. *Our experience is that a combination of matched and mismatched strains may circulate together and often at different times during the season. The vaccine strain is thus matched to a proportion of circulating strains, ranging from zero to one. Match/mismatch is not an all or nothing phenomenon. How have the authors dealt with this?*

Thank you for pointing this out. We characterized breakthrough strains and assumed equal exposure of participants in both the vaccine(s) and control arms. This was conducted for mismatched influenza A strains (lineage drift) and mismatched influenza B strains (lineage mismatch and mismatch by antigenic drift). Most of the studies report on whether the strains were a match or mismatch using a cut-off of 4-fold increase/decrease in HI titres and they don’t report the specific degree of mismatch (i.e., most studies dichotomize match/mismatch). We agree that this is a limitation and have noted this in our discussion section (see lines 393-396): “For the purposes of our analysis, we dichotomized cross-protection but in reality, the degree of mismatch is a continuum. The cross-protection inferred by mismatch strains should be analyzed as a continuum in the future.”

d. *Match may be good for one or more of the vaccine strains and poor for one or more. When reporting overall VE against mismatched strains, how have you decided the ‘overall vaccine’ (all 3 strains) was mismatched?*

We only included influenza infections that were due to mismatched strains in the mismatched analyses. Similarly, we only included influenza infections that were considered matched in the matched analysis. Therefore, we are not determining the overall vaccine efficacy of all three strains. To clarify, we have added the following text in our synthesis section (lines 179-181) “Only influenza infections due to mismatched strains were included in the mismatched analysis,
while only influenza infections due to matched strains were included in the matched analysis.”

We have also moved important details regarding the ascertainment of match/mismatch in Table 1 of our manuscript (this information was originally hidden in Appendix 2), as per the request of the second reviewer.

6. Although you claim that VE against matched strains is higher than VE against mismatched strains, differences in VE have not been formally tested. In some instances, the general claim does not hold. For instance, for TIV, influenza A and adults, VE against matched strains was 64% (52,73) whereas VE against mismatched strains was 61% (9,84) (p15). These estimates are not likely to differ on formal testing. Moreover, even when the point estimate for matched strains was higher than the point estimate for mismatched strains, confidence intervals sometimes overlapped.

In our discussion and abstract, we have now added that the overall VE is “slightly” higher for matched versus mismatched strains.

7. The use of serology is not optimal as an endpoint for assessing VE against TIV, giving potentially anomalous results, as described here. In this context it might be worth noting the work of Monto’s group on the choice of endpoints: JID 2011; 203:1309-15.

We completely agree and have added this reference to line 401 of the discussion.

8. Another apparently anomalous finding was of a higher point estimate for protection against unmatched strains for LAIV in adults. The authors attribute this to possible confounding (p17). Is this strictly a confounder, as confounding is formally understood?

We agree that our use of the term confounding is not appropriate and have deleted it. Our explanation is simply due to a possible discrepancy in the degree of matching. As noted above, we didn’t look at the degree of drift/distance of the drift and considered this a dichotomous versus a continuous outcome, which could potentially have impacted our vaccine efficacy results. This has been revised in the discussion (lines 389-396): “This finding might suggest that there might be a possible discrepancy in the degree of matching, which may have impacted our results. Specifically, that trials conducted amongst children may have had a greater degree of mismatch than those conducted amongst adults. Unfortunately, the current analysis does not allow the degree of mismatch to be examined. For the purposes of our analysis, we dichotomized cross-protection but in reality, the degree of mismatch is a continuum. The cross-protection inferred by mismatch strains should be analyzed as a continuum in the future.”

9. I think the authors need to provide further discussion on possible explanations for their findings. Can part of the findings be explained by some residual misclassification of matching, as explored in comment 5, above? Some of the findings are no doubt related to the poor recent discriminatory ability of H1 assays for H3 subtypes. This deserves further exploration. Finally, significant cross protection that is not detected by HI assays may exist, and this could be explored.
We agree fully with the reviewer that some of the findings may have been explained by residual misclassification of matching due to the limited discriminatory ability of H1 assays. We have added the following text to the discussion section (lines 427-437): “A critical limitation, which may have influenced our results, is the determination of mismatch between circulating strains and those found in the vaccine. Characterizing strains as antigenically similar or distinct using HI assay or ferret antisera might be insensitive, leading to misclassification of strains [PMID: 15218094]. Residual misclassification of matching due to the limited discriminatory ability of H1 assays may have also explained some of our findings and is a limitation of the inferences that can be made. Furthermore, the cut-off values recommended by the CDC to distinguish between match and mismatch strains using the HI assay changed during the study period across the included RCTs, although most of the studies included here used a 4-fold quotient HI cut-off (Table 1). As such, some of the data from trials labeled as matched might actually have been mismatched [59], and vice versa.”

10. In a pooled 5 year observational study, we found similar results to those reported here, specifically VE for TIV in adults 20-64 years of 62% and no obvious association of VE with an assessment of match (IRV published online 17 October 2012 doi: 10.111/irv.12018 ). The results from this field broadly study support the conclusions of the meta-analysis of trials.

We thank the reviewer for this reference. We have included it in our discussion section (lines 382-384): “Our results are consistent with those found in a pooled observational study including 5 years of data [PMID: 23078073].”

REVIEWER 2: Walter Emil Philipp E Beyer
see also attachment: 2700060098655244_article_remarks.doc

1. The Results section should be completely rewritten in a way more accessible to the reader. It is not necessary to present every single confidence interval in the text. Use tables and figures for the details and focus on the main results in the text.

We have significantly reduced the results, by focusing on the LAIV in children and the TIV in adults for the primary outcome only.

Minor remarks and comments are directly written into the manuscript, after conversion of the pfd-file to a Word-file. See document2700060098655244_article_remarks.doc.

Thank you for your diligence. We have made most of your suggested changes to the text.

2. Major comments:

1. Classification of vaccine types and formulations

The present classification (LAIV, TIV other) is not valid.

Two relevant vaccine types are to be covered:
(1) The systemically administered (im, sc) inactivated influenza vaccine (IIV), including different formulations: whole-virus, split (subvirion) and subunit, adjuvanted or not.

(2) The intranasally administered live influenza vaccine (LIV).

The valence of the vaccines (how many strains included) is not relevant for this review. Suppose, the vaccine contains strain XY. Then the only relevant questions for this review are whether the epidemic strain is reported, whether it belongs to the same (sub)type as XY, and if yes, whether it matches XY or not. It is not relevant whether the vaccine, besides XY, contains other influenza strains or not. Thus, mono-, bi-, tri- and tetravalent vaccines should all be classified as either IIV or LIV. It makes no sense (not to say: it is misleading) to treat a trivalent vaccine differently from a bivalent vaccine. Valence is accidental and reflects only the epidemiological situation. Up to 1978, vaccines were usually bivalent, as only A-H3N2 and B circulated. Only after the re-occurrence of A-H1N1, the vaccine became trivalent.

In three RCTs, recombinant hemagglutinin vaccine was used. This is simply an inactivated vaccine (comparable to the subunit formulation) and should be included in that stratum. Thus, vaccines should be categorized as follows:

Ref. First author Current manuscript Revised strata
24 Leibovitz other MIV IIV
25 Beutner other MIV IIV
26 Rytel other bivalent LAIV LIV
27 Monto other bivalent LAIV LIV
28 Tannock TIV IIV
29 Keitel Y1 other WV IIV
29 Keitel Y2 other WV IIV
29 Keitel Y3 other WV IIV
29 Keitel Y4 other WV IIV
29 Keitel Y5 other WV + split IIV
30 Gruber TIV IIV
31 Edwards A LAIV LIV
31 Edwards B other BIV IIV
32 Clover A TIV IIV
32 Clover B other bivalent LAIV LIV
33 Govaert other QIV IIV
34 Powers A TIV IIV
34 Powers B other rHA IIV
35 Belshe LAIV LIV
36 Rudenko A LAIV LIV
36 Rudenko B TIV IIV
37 Belshe LAIV LIV
38 Bridges TIV IIV
39 Hoberman TIV IIV
40 Tam LAIV LIV
41 Vesikari LAIV LIV
42 Forrest LAIV LIV
12 Bracco Net LAIV LIV
43 Lum LAIV LIV
All trials can now be included in a hierarchic structure:
Vaccine type # Age class # (Primary/secondary outcome) # Match. This should also be the sequence for the Results section.

Thank you for your comment and for going through the studies for us – we greatly appreciate your time and effort on this. The planned analysis of separating the vaccines by TIV, LAIV and other was carefully thought through to reflect the current influenza vaccine climate in most countries. We have published these methods in a peer-reviewed journal (Sys Rev) in our systematic review protocol, thus are quite hesitant to change the way that we’ve categorized these studies. However, we do see the validity in your comment and have revised the methods to state that we conducted a sensitivity analysis categorizing the studies in the manner that you suggest above (lines 176-177): “A post hoc sensitivity analysis was also conducted to examine the influence of categorizing the vaccines as being either an inactivated influenza vaccine or a live influenza vaccine.”

As you can see in the table below, the results achieved via this categorization are virtually identical to the results we present in our paper. And as reviewer 3 notes, by excluding vaccines with non-standard dosages and modes of delivery, we are actually presenting the results of a cleaner analysis. As such, we have reported the following in our results (lines 356-359): “Post hoc sub-group analysis: Our results did not change after a post hoc sensitivity analysis was conducted to examine the influence of categorizing the vaccines as being either an inactivated influenza vaccine or a live influenza vaccine.”

<table>
<thead>
<tr>
<th>Category</th>
<th># influenza seasons</th>
<th># patients</th>
<th># influenza cases (vaccine group)</th>
<th># influenza cases (control group)</th>
<th>I² (%)</th>
<th>VE (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIV matched</td>
<td>15</td>
<td>28114</td>
<td>255</td>
<td>844</td>
<td>85%</td>
<td>0.77 (0.66, 0.85)</td>
</tr>
<tr>
<td>LAIV matched</td>
<td>12</td>
<td>28114</td>
<td>255</td>
<td>844</td>
<td>85%</td>
<td>0.77 (0.66, 0.85)</td>
</tr>
<tr>
<td>LIV mismatched</td>
<td>16</td>
<td>27241</td>
<td>219</td>
<td>424</td>
<td>65%</td>
<td>0.61 (0.46, 0.71)</td>
</tr>
</tbody>
</table>
3. Table 1 in the present form contains much irrelevant information. Nobody is really interested in the fact that the placebo formulation in Belshe 2000 contained sucrose, phosphate and glutamate, for example. The column ‘Placebo composition’ can be removed, as well as the following ‘Route of administration’ as this information is already inherent to the vaccine type. Importantly, the influenza B vaccine strains should all be characterised as either Yamagata or Victoria lineage.

We agree and have moved all of the irrelevant information to the appendix and have added the important information (including the characterization of influenza B) in Table 1. As you can see, Table 1 now includes the following columns:
- Author and year
- Vaccine type (TIV, LAIV, other)
- Vaccine composition
- Type of lab confirmed influenza used in the analysis
- Circulating strains
- Antigenic characterization as per author (including the HI titre cut-off values)
- Classification of lab-confirmed influenza viral strains as being matched
- Classification of lab-confirmed influenza viral strains as being mismatched

4. Table 2
Edwards 1994: “All groups” under Age category should read: “Children/adults” as no elderly persons were included.

As suggested, we have made this change to Table 2.

5. Presentation of antigenic match
This essential information, in fact the core of the whole work, is now hidden in Appendix 2, but should prominently appear in the main text, freed from irrelevant details and possibly linked to Table 1. Taxonomic names of virus strains should be uniform throughout the table and according to WHO style (i.e., not ‘A/Port Chalmers’ or ‘A/Port Chalmers (H3N2)’, as in Beutner 1979, but ‘A/Port Chalmers/1/73 (H3N2)’).

As suggested, we have added this important information to Table 1.

Beutner 1979: ‘A/Victoria (H1N2)? That should be an A-H3N2 strain.

This has been revised, as suggested.

Clover 1991 is classified as ‘match’, but is actually a major mismatch.
This has been revised Table 1 and in our secondary analysis (this study only provided data on our serologic assay outcome).

REVIEWER 3: Nancy Cox

1. **Major Compulsory Revision:** The authors reviewed clinical trials that measured influenza vaccine efficacy (VE) and stratify results by whether circulating viruses “matched” the vaccine viruses or were "mismatched"/"unmatched". While this is a very thorough review and the meta-analysis methods appear solid, it is essential to improve the manuscript by defining what criteria were used to define "matched" vs. "unmatched" or "mismatched" vaccine viruses. The authors leave the readers unclear what the definition for matched and unmatched was, e.g. the cut off for HI that determined whether a virus was antigenically similar to a vaccine virus or not; cut off values were likely different depending on the study and/or year of the study. While, the authors discuss this issue in their limitations it is essential to understand the definition for “mismatched” for each study. Otherwise, no other researchers can replicate the work in an independent fashion.

Thank you for pointing this out. We have now inserted the cut-off values for HI for all of the included studies in Table 1. You will see that most used an HI quotient of ≥4-fold titre to identify mismatched strains and an HI typing quotient of <4-fold titre to identify matched strains. The exceptions here are Beutner 1979 (but this was classified as an “other vaccine” so was not in our main analysis), Clover 1991 (but this was included in our secondary analysis because influenza infection was not determined via PCR or culture so this does not impact our main results), Monto 2009 (when we removed this in our sub-group analysis our results did not change) and Barrett 2011 (when we removed this in our sub-group analysis our results did not change).

2. Furthermore, clinical trial factors are likely related to VE, including age and prior exposure to influenza viruses in the past. This may be reflected in the large difference noted in VE from matched and mismatched years among children but the small differences reported for adults (e.g. TIV 64% and 61%, respectively). In addition, age may be an effect modifier so lumping young children, who have a low likelihood of prior exposure, together with older individuals makes little sense when you are comparing VE for matched and mismatched vaccines. Finally, the study, treats all RTCs equally and some of the trials were conducted with non-standardized doses and routes of vaccination. Using stricter criteria for choosing which RTCs to include would improve the quality of the results (see below).

We have run a sub-group analysis on children only and did indeed wish to further break this down as older children and younger children. However, as noted in the table, all of the studies overlapped the age category of 6-36 months. As such, we were unable to tease this out further, unfortunately.

For LAIV matched we have 9 influenza seasons

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Institute</th>
<th>Vaccine</th>
<th>Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>366 Y1</td>
<td>Primary Tam</td>
<td>2007</td>
<td>Children (12-36 mos)</td>
<td></td>
</tr>
<tr>
<td>411 Y1</td>
<td>Primary Vesikari</td>
<td>2006</td>
<td>Children (6-36 mos)</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Type</td>
<td>Vaccine</td>
<td>Year</td>
<td>Age Group</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>------</td>
<td>-----------</td>
</tr>
<tr>
<td>411 Y2</td>
<td>Prim ary</td>
<td>Vesikari</td>
<td>2006</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>844 Y1</td>
<td>Prim ary</td>
<td>Bracco Neto</td>
<td>2009</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>844 Y2</td>
<td>Prim ary</td>
<td>Bracco Neto</td>
<td>2009</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>1001</td>
<td>Prim ary</td>
<td>Belhine</td>
<td>2000</td>
<td>Children (6-23 mos)</td>
</tr>
<tr>
<td>312</td>
<td>Prim ary</td>
<td>Forrest</td>
<td>2008</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>1002</td>
<td>Prim ary</td>
<td>Belhine</td>
<td>1998</td>
<td>Children (6m-1.5 yrs)</td>
</tr>
<tr>
<td>1003</td>
<td>Prim ary</td>
<td>Lum</td>
<td>2010</td>
<td>Children (11-24 mos)</td>
</tr>
</tbody>
</table>

For LAIV m.im matched we also have 9 influenza seasons

<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Vaccine</th>
<th>Year</th>
<th>Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>366 Y1</td>
<td>Prim ary</td>
<td>Tam</td>
<td>2007</td>
<td>Children (12-36 mos)</td>
</tr>
<tr>
<td>366 Y2</td>
<td>Prim ary</td>
<td>Tam</td>
<td>2007</td>
<td>Children (12-36 mos)</td>
</tr>
<tr>
<td>411 Y1</td>
<td>Prim ary</td>
<td>Vesikari</td>
<td>2006</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>411 Y2</td>
<td>Prim ary</td>
<td>Vesikari</td>
<td>2006</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>844 Y1</td>
<td>Prim ary</td>
<td>Bracco Neto</td>
<td>2009</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>844 Y2</td>
<td>Prim ary</td>
<td>Bracco Neto</td>
<td>2009</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>1001</td>
<td>Prim ary</td>
<td>Belhine</td>
<td>2000</td>
<td>Children (6-23 mos)</td>
</tr>
<tr>
<td>312</td>
<td>Prim ary</td>
<td>Forrest</td>
<td>2008</td>
<td>Children (6-36 mos)</td>
</tr>
<tr>
<td>1003</td>
<td>Prim ary</td>
<td>Lum</td>
<td>2010</td>
<td>Children (11-24 mos)</td>
</tr>
</tbody>
</table>

Regarding the inclusion of non-standardized doses and routes of vaccination, most of these studies were classified as being under the “other vaccine” category so were not included in our analysis of the TIV and LAIV. The only exceptions were Tannock 1984 and Rudenko. However, these studies did not report influenza infection as per culture or PCR so were not included in our primary outcome analysis, having no impact on our main findings.

**Minor essential revisions:**

3. Intro – reference #2 should be replaced with more recent estimates (MMWR)

Reference 2 has been deleted in the first sentence, as recommended by reviewer 1.

2. Figures – Figure legends appear to be missing.

These have been added, as suggested.
3. Methods: The authors should specify whether they included efficacy estimates for young children who had received one or two doses of vaccine.

The dosages have been reported for each included RCT in Appendix 2.

4. Methods - included studies: All vaccines appear to be considered equally. Yet, there is variability in the antigen dose of the vaccines that could influence VE. Older studies using CCA units/ml may have imprecise doses. Can the authors discuss? Also, the route of delivery might be important, for example intranasal spray versus intranasal drops and IM versus SC.

Please see our response to comment 2 above.

5. Results: page 13 – Can the authors discuss their findings for LAIV in adults where “mismatched” vaccines appeared to have higher VE estimates than “matched” vaccines for adults.

We have brought this up in our discussion section (lines 385-396), as follows “We found that the LAIV was more efficacious among children versus adults, which is likely a reflection of the difference in previously acquired influenza infections between age groups and the consequently larger amount of pre-vaccination antibody, which affects the live vaccine. However, we found higher efficacy for adults versus children for mismatched LAIV estimates. This finding might suggest that there might be a possible discrepancy in the degree of matching, which may have impacted our results. Specifically, trials conducted amongst children may have had a greater degree of mismatch than those conducted amongst adults. Unfortunately, the current analysis does not allow the degree of mismatch to be examined. For the purposes of our analysis, we dichotomized cross-protection but in reality, the degree of mismatch is a continuum. The cross-protection inferred by mismatch strains should be analyzed as a continuum in the future.”

6. Discussion - page 17, lines 327-334 - Without any data or report to substantiate a “degree of drift” argument, this is a less than satisfactory argument. Can you elaborate and provide references that substantiate this argument for the years these studies took place. Also, this paragraph oversimplifies what VE measures and how a virus is characterized to determine antigenic similarities. In reality, genetic and extensive serologic data supplement HI data when determining antigenic similarities with vaccine viruses.

We have revised this paragraph (see response to comment 5 above) and hope that it is satisfactory now. As noted in our response to comment 1 above, the cut-offs used for HI data were consistent across the majority of the studies included in our main analyses.

7. Discussion page 19 (line 364). Potential misclassification of the major outcome of the study is a major limitation. There is also inter-laboratory variability of HI testing. Therefore, it would be very helpful to know what cut off values were used for each study and what the observed HI value was for the viruses in these studies.

We agree and please see our response to comment 1 above.