PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Association between rhinovirus wheezing illness and the development of childhood asthma: a meta-analysis
AUTHORS	Liu, Lu; Pan, Yilin; Zhu, Yanting; Song, Yang; Su, Xiaofan; Yang, Lan; Li, Manxiang

VERSION 1 - REVIEW

REVIEWER	Scott Montgomery Örebro University Sweden
REVIEW RETURNED	07-Jul-2016

GENERAL COMMENTS	The paper by Liu and colleagues investigates whether rhinovirus infection with contemporaneous wheezing illness is associated with subsequent wheeze or asthma. Below are some comments. 1. The number of relevant papers identified is small. The search strategy required presence of the terms ‘rhinovirus’ or ‘virus’. Would use of the alternative terms ‘infection’ or ‘infections’ have identified a larger number of relevant publications? 2. The formatting of table 1 should be revisited as it is rather difficult to read at the moment. 3. The authors conclude that the association of early infection with wheeze with later wheeze or asthma does not persist beyond age 10 years. While I am inclined to believe this as it is consistent with other studies (infections in infancy and later asthma), the conclusion should be made more cautiously as it is based on only two studies. 4. Differences in aetiology and the role of atopy should be discussed more fully as, for example in the study by Kusel et al, early infection is associated with later asthma among those who are atopic after age two years. 5. How reliable is the identification of rhinovirus across the studies and to what extent is it likely to be misclassified (not identified), and could this influence the results? 6. The causes of wheezing and asthma in children (the outcomes of this study) are not always aetiological identical, so the consequences of this heterogeneity in outcome should be discussed.

REVIEWER	JANVIER GASANA Kuwait University

	Jabriya Kuwait
REVIEW RETURNED	10-Jul-2016

GENERAL COMMENTS	The authors have to provide an exhaustive literature review that describes and agrees with their findings. What are the pathophysiological mechanisms of the association of RV and the subsequent development of wheezing/asthma? The authors need to provide a valid explanation of why they included studies by Hyvarinen, Kusel, Gupten, and Teeratakulpsarn that have the lower limit of the confidence interval below 1. PRISMA is attached to the manuscript while it is not mentioned anywhere in the text of the manuscript.

REVIEWER	Ryan Pohlig University of Delaware, United States of America
REVIEW RETURNED	04-Oct-2016

GENERAL COMMENTS	Methods: Can the authors better specify the boolean searchers used? I think what they meant is that initial measures were as infants and follow up at adolescence, but using (infants or young children) and (children or childhood or adolescence) seems it would require an article to have both phrases compared to an article the word children twice. Authors should considering running a second analysis using a common follow-up time and not just last one reported and include results if different or mention if same. Was sample size not collected as part of criteria? Importantly, authors should also considered non-categorical outcomes, and converted them to either an OR/HR/RR or converted all effect sizes into a standardized form (i.e. Hedge's). Almost 400 articles reduced to 9 is a bit drastic. This is a big power concern for analyses. Please cite a statistical source for assuming that OR=HR=RR, not another meta-analysis/systematic review. There were 0 studies before 2005 that examined this issue? Restults: Funnel plot was reported as used but not depicted. In a funnel plot either sample size or estimate of variance is used on the Y axis (this could also help with potential publication bias identification). It is recommended a true funnel plot be included before publication. Why is subgroup analysis performed by race? Adds nothing to paper as results are consistent with main findings (same direction), also having only 2 studies in a subgroup should not result in any model tests of significance- that I am aware of, please provide a citation for one.

	What strata were used for stratificatino of age at follow up? In the table only 2 articles have follow up 10 or more years and the association was not found there. Discussion: p9 line 56, this comment does not make sense in light of findings. You found studies less than 10 years were significant, which would go against more exposure as the follow up times would be shorter than the 10+ studies. p10 line 24, meta-analysis does not create a larger sample size, in-fact yours (n=9) was smaller than each study. minor comments Methods: "Besides" can be deleted p7 line9, this approach can be called jackknifing. Stop using the ">" sign for subgroup identification. This might help eliminate the error made in discussion.

REVIEWER	Amalia Karahalios University of Melbourne, Australia
REVIEW RETURNED	05-Oct-2016

GENERAL COMMENTS	This is a well written meta-analysis looking at the association between rhinovirus wheezing illness and childhood asthma from 9 cohort studies. I have some major and minor comments on the manuscript which are detailed in my review (see below). This is a reasonably well-written meta-analysis of the association between the rhinovirus wheezing illness and childhood asthma from 9 cohort studies. The results of the fixed-effect meta-analysis showed an increased odds of wheezing/asthma associated with rhinovirus illness in the first 3 years of life. The strengths of the study include its primacy, and use of cohort studies. However, there are a few specific comments and suggestions summarized below. **Major comments** - Did the authors register their review with PROSPERO http://www.crd.york.ac.uk/PROSPERO/ or another similar database prior to commencing their review? Similar to RCTs, a research protocol should be drafted and published for systematic reviews and meta-analyses to ensure scientific rigor. - The search strategy is not detailed enough - for example in PubMed you should be using MeSH terms where appropriate and similar for the other databases. It is necessary to undertake a comprehensive search of the literature using an appropriate search strategy for each database searched. This ensures that the search is in fact systematic, reproducible and identifies all of the relevant literature. A complete search for at least one of the databases should be included as supplementary material. I suggest speaking to a librarian or a similarly qualified

	individual. Further, a qualification of the search strategy should be provided (if undertaken), for example did the authors consult librarians? - How did you deal with articles presenting results from the same study? I assume that this is what 'redundant information' pertains to in figure 1. - Use of Newcastle-Ottawa Scale is controversial as the summary scores involve inherent weight of component items (Greenland 1994 “Quality scores are useless and potentially misleading” American Journal of Epidemiology, 300-301); the authors should discuss the limitations of these scores in their discussion. - In the methods section the authors state that they assumed the HR and RR approximated the OR. This is only true if the outcome is rare. Is asthma considered to be a rare outcome (I wouldn’t think so)? If not, this is not appropriate. - The authors rely on the use of “statistical significance”; I strongly suggest that the authors read the recent paper by Greenland et al “Statistical tests, P values, confidence intervals and power: a guide to interpretations”. European Journal of Epidemiology. 2016; 31:337-350 and that they remove any references to “statistical significance”. - The authors state that heterogeneity was confirmed from a p-value of 0.1 or I^2 of more than 50%; how was this chosen? - The authors indicate that they carried out a sensitivity analysis by sequentially excluding each study; a reference is needed for this type of sensitivity analysis. This type of sensitivity analysis is not recommended; studies should be excluded for reasons chosen a priori. - More information is needed in the Table 1. I suggest including the number of participants in each cohort, the number of participants with asthma, the follow-up time, the statistical method used, the confounding variables adjusted for in the analysis, and the amount of missing data. If necessary, some of this information could be provided in a supplementary table. - In the first paragraph of the discussion, the authors mention confounding. The main problem with cohort studies is confounding and I don’t feel that this issue has been appropriately handled in this manuscript. A further discussion of potential confounding is needed, as is a deeper exploration of how each individual study has dealt with confounding in their analysis. Please refer to Egger et al “Spurious precision? Meta-analysis of observational studies”. BMJ. 1998 pages 140-144 for further reading. - In the discussion the authors mention that the diagnosis of

wheezing/asthma was not always provided by a physician; I wonder if this would create a bias and whether it is worth carrying out a subgroup analysis of papers with/without physician diagnosis of wheezing.

- The authors state that there is no evidence of publication bias based on funnel plot and Egger's test; however Egger's test is known to have low power for meta-analysis of fewer than 20-25 studies. This meta-analysis includes only 9 studies; the limitations of this test should be discussed.

- In light of my previous comments about confounding, the last line of the conclusion needs to be refined and made more specific.

- In figure 1, it is not clear to me why 3 studies "with no subgroup analysis" were excluded. Can't these studies be included in the overall analysis?

Minor details

Abstract

1. More details of the methods used to synthesise the data (i.e. the meta-analysis methods) should be included in the abstract.
2. It is not necessary to include the name of the software package in the abstract.
3. Reword the sentence on line 34 "There was a significant…" to "RV wheezing illness in the first 3 years of lie was associated with an increased odds of wheezing/asthma in later life (OR = 2.02; 95% CI = 1.58 – 2.57, P<0.001)."
4. Reword the end of the sentence on line 43 from "…was only observed among < 10 years studies" to "was only observed for studies that had less than 10 years of follow-up".
5. The last sentence of the conclusion "Larger-scale and well-designed studies are…" is too vague.

Introduction

1. Page 4 Line 30 – RV-related diseases "are" globally…
2. Page 4 Line 38 – reference needed for sentence ending with "… for virus replication".
3. Page 5 line 10 – reword "no study has been published which has combined systematic review and synthesis of the evidence." to "no study has been published that has systematically reviewed the literature and synthesized the available evidence".
4. Page 5 line 12/13 – we performed "a systematic review and" meta-analysis…

Methods

1. Indicate the start date of the search no just the end date.

	2. Page 6 line 31/32 – "outcome measure" should be changed to "effect estimate" 3. Page 6 line 52 – What do you mean by "by using data reported in the relevant literature". Please clarify and include appropriate references. 4. Page 7 line 9 – include categorisations for ethnicity and age at follow-up **Results** 1. Table 1 – present age in the same unit for every study (e.g. months) 2. Table 1 is difficult to read. I suggest putting it in landscape format. 3. Page 9 line 24/25 – change the wording "… was only observed among < 10 years studies…" to "… was only observed among studies that had a follow-up time of less than 10 years…". I would also suggest presenting the results for studies with follow-up time > 10 years in this paragraph. 4. Page 9 line 39/40 – include the funnel plot as a figure 5. Page 9 line 58 – change the wording "… was only observed among <10 years studies." To "… was only observed among studies that had a follow-up time of less than 10 years". **Discussion** 1. Page 10 line 22 – sentence ending with "…development of subsequent wheezing and asthma." needs a reference. **Figures** 1. It is helpful to show how many records were identified in each database and how many duplicate records were excluded. **Supplementary material** 1. Item #5 in the PRISMA checklist should be marked as 'Not reported'

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:
Comment 1: The number of relevant papers identified is small. The search strategy required presence of the terms 'rhinovirus' or 'virus'. Would use of the alternative terms 'infection' or 'infections' have identified a larger number of relevant publications?
Response 1: We thank the reviewer for this helpful suggestion. We added "infection" to the search terms and re-searched the databases. The number of studies identified was increased to 7082, and 15 articles containing the same cohorts with different follow-up periods met our eligibility criteria and were included in this meta-analysis. We have modified the methods and results sections in the revised manuscript (Page 5, lines 26-30; Page 7, lines 35-40).

Comment 2: The formatting of table 1 should be revisited as it is rather difficult to read at the moment.
Response 2: We are sorry for this issue. We have revised the table 1 and put it in landscape format (Page 8).

Comment 3: The authors conclude that the association of early infection with wheeze with later wheeze or asthma does not persist beyond age 10 years. While I am inclined to believe this as it is consistent with other studies (infections in infancy and later asthma), the conclusion should be made more cautiously as it is based on only two studies.
Response 3: We agree with the reviewer's point that the conclusion should be made more cautiously. We have included all relevant studies (n=15) containing the same cohorts with different follow-up periods to perform subgroup analysis stratified by age at follow-up, and the results indicated that the association still remained significant in both less than 10 years (11 studies) and more than or equal to 10 years (4 studies). In light of the potential confounding factors and limitations, the results should be interpreted more cautiously and more large-size prospective studies are required to validate the risk identified in the current meta-analysis. We have revised the manuscript according to our results and your suggestion (Page 9, lines 16-25; Page 13, lines 10-15).

Comment 4: Differences in aetiology and the role of atopy should be discussed more fully as, for example in the study by Kusel et al, early infection is associated with later asthma among those who are atopic after age two years.
Response 4: Asthma is a multi-factorial disease resulting from complex interactions between genetic predisposition and environmental factors. Although the results of some studies included in our meta-analysis were adjusted by factors which might alter the association between RV wheezing and wheezing/asthma risk, there were still other potential influences factors. We have discussed this in the revised manuscript (Page 12, lines 31-38). Atopy has been defined as a pivotal risk factor for both RV wheezing illness and development of asthma, which might be a confounding factor or interact with RV infection. We have added the discussion about the role of atopy in the revised manuscript (Page 10, lines 25-35).

Comment 5: How reliable is the identification of rhinovirus across the studies and to what extent is it likely to be misclassified (not identified), and could this influence the results?
Response 5: We thank the reviewer for inquiring this. All of the studies included in our meta-analysis defined RV by PCR. PCR is a quite sensitive approach to detect virus infection, it should not miss any RV infection. However, the high sensitivity of PCR might cause false positive results, as the presence of virus nucleic acid in respiratory secretions of patients with respiratory symptoms does not prove that the virus is the cause of the symptoms. We have discussed this in the revised manuscript (Page 12, lines 8-18).

Comment 6: The causes of wheezing and asthma in children (the outcomes of this study) are not always aetiological identical, so the consequences of this heterogeneity in outcome should be discussed.
Response 6: We agree with the reviewer that the causes of wheezing are not always identical to that of asthma. However, the children with wheezing in the earlier stage of life will easily suffer from asthma, and some children with early onset asthma will get better with time, so it's difficult to distinguish between wheezing and asthma. We have discussed this heterogeneity in outcome in the revised manuscript (Page12, lines 24-31).

Reviewer 2:
Comment 1: The authors have to provide an exhaustive literature review that describes and agrees with their findings.
Response 1: We thank the reviewer for this helpful suggestion. We have thoroughly reviewed the literature and made a discussion to support our findings in the revised manuscript (Page 9, lines 45-

48; Reference 12, 24, and 40).

Comment 2: What are the pathophysiological mechanisms of the association of RV and the subsequent development of wheezing/asthma?
Response 2: We have discussed the pathophysiological mechanisms of the association of RV and the subsequent development of wheezing/asthma in the discussion part of the revised manuscript (Page 11, lines 6-20).

Comment 3: The authors need to provide a valid explanation of why they included studies by Hyvarinen, Kusel, Gupten, and Teeratakulpsarn that have the lower limit of the confidence interval below 1.
Response 3: We thank the reviewer for inquiring this. The lower limit of the confidence interval below 1 indicates no statistical difference in the outcome between exposure and non-exposure subjects. Results of these studies are inconsistent with other studies which indicate RV can increase the risk of asthma. We have included all relevant studies with different results in order to synthesize the available evidence and make the conclusion more cautious.

Comment 4: PRISMA is attached to the manuscript while it is not mentioned anywhere in the text of the manuscript.
Response 4: This study was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We have added this in the revised manuscript (Page 5, lines 13-15).

Reviewer 3:
Comment 1: Can the authors better specify the boolean searchers used? I think what they meant is that initial measures were as infants and follow up at adolescence, but using (infants or young children) and (children or childhood or adolescence) seems it would require an article to have both phrases compared to an article the word children twice.
Response 1: We thank the reviewer for this helpful suggestion. We have revised the search terms and re-searched the databases, more details have been provided in the revised manuscript (Page 5, lines 26-30).

Comment 2: Authors should consider running a second analysis using a common follow-up time and not just last one reported and include results if different or mention if same.
Response 2: We thank the reviewer for this helpful suggestion. We have included all relevant studies (n=15) containing the same cohorts with different follow-up periods to perform subgroup analysis stratified by age at follow-up and provided the results in the revised manuscript (Page 9, lines 16-25).

Comment 3: Was sample size not collected as part of criteria?
Response 3: We did not define sample size as a part of criteria in order to include more relevant studies and comprehensively evaluate the association.

Comment 4: Importantly, authors should also considered non-categorical outcomes, and converted them to either an OR/HR/RR or converted all effect sizes into a standardized form (i.e. Hedge's).
Response 4: We thank the reviewer for this suggestion. The outcomes in studies included in our meta-analysis were all categorical variables.

Comment 5: Almost 400 articles reduced to 9 is a bit drastic. This is a big power concern for analyses.
Response 5: We have checked the studies according to the inclusion and exclusion criteria and excluded duplicates and studies that did not report on RV wheezing illness and wheezing/asthma outcome or did not provide available data. More details of study selection have been described in the revised manuscript (Page 7, lines 35-40).

Comment 6: Please cite a statistical source for assuming that OR=HR=RR, not another meta-analysis/systematic review.
Response 6: We assumed that OR=HR=RR according to the study published previously (Greenland S, Epidemiol Rev, 1987). While this is only true if the outcome is rare, and asthma might not be considered as a rare outcome, so this assumption is inappropriate. By reviewing the relevant literature, we used RRs as the common measure of association across studies, directly considered HRs as RRs and transformed ORs into RRs using the formula RR=OR/ [(1-P0) (P0×OR)] (Zhang J, et al, JAMA, 1998). We have modified the methods sections in the revised manuscript (Page 6, lines 48-53).

Comment 7: There were 0 studies before 2005 that examined this issue?
Response 7: Kotaniemi-Syrjänen, et al. have published their study in 2003 (Kotaniemi-Syrjänen, et al. J Allergy Clin Immunol, 2003), while Hyvärinen, et al. from same group have reported the results of a same longitudinal cohort with a longer follow-up period (Hyvarinen MK, et al. Pediatric pulmonology, 2005). According to the original inclusion criteria, Hyvärinen study with a longer follow up period was only incuded, and Kotaniemi-Syrjänen study with a shorter follow up period was excluded. We have modified the inclusion criteria and included all relevant studies (n=15) containing the same cohorts with different follow-up periods to perform subgroup analysis stratified by age at follow-up according to the suggestion of reviewers. The details of included studies have been shown in table 1 of revised manuscript (Page 8).

Comment 8: Funnel plot was reported as used but not depicted. In a funnel plot either sample size or estimate of variance is used on the Y axis (this could also help with potential publication bias identification).
Response 8: We thank the reviewer for this helpful suggestion. We have depicted the funnel plot in the revised manuscript (Page 17, lines 15-22).

Comment 9: It is recommended a true funnel plot be included before publication.
Response 9: We have provided the funnel plot as figure 5 in the revised manuscript (Page 9, lines 36).

Comment 10: Why is subgroup analysis performed by race? Adds nothing to paper as results are consistent with main findings (same direction), also having only 2 studies in a subgroup should not result in any model tests of significance- that I am aware of, please provide a citation for one.
Response 10: It commonly thought that asthma is a multi-factorial disease resulting from complex interactions between genetic predisposition and environmental factors. We performed subgroup analysis stratified by race to evaluate whether different genetic backgrounds influenced the association between RV wheezing illness in the early life and subsequent development of wheezing/asthma. Although the results remained significant, it needed further validation because there were only 2 studies in Asians. We have discussed this in the revised manuscript (Page 9, lines 53-57; Page 10, lines 3-8). Some meta-analyses also have only 2 studies in a subgroup when performing subgroup analyses (Gijsbers L, et al. Nutr Metab Cardiovasc Dis, 2016; Xiang Z et al. Oncotarget, 2016; Glazov G, et al. Acupunct Med, 2016). We will carry out the analysis again to obtain a more reliable result when the number of relevant studies increases.

Comment 11: What strata were used for stratification of age at follow up? In the table only 2 articles have follow up 10 or more years and the association was not found there.
Response 11: It has been shown the association between respiratory syncytial virus hospitalization and wheezing/asthma decreases with age at follow-up (Régnier SA, et al. Pediatr Infect Dis J, 2013). To evaluate whether the association between RV wheezing and development of wheezing/asthma would change with age, here we performed similar analysis. We have included all relevant studies

(n=15) containing the same cohorts with different follow-up periods to perform subgroup analysis stratified by age at follow-up, and the results indicated that the association still remained significant in both less than 10 years (11 studies) and more than or equal to 10 years (4 studies). We have modified the methods and results sections in the revised manuscript (Page 7, lines 18-22; Page 9, lines 16-25).

Comment 12: p9 line 56, this comment does not make sense in light of findings. You found studies less than 10 years were significant, which would go against more exposure as the follow up times would be shorter than the 10+ studies.
Response 12: There is a study indicated that the association between respiratory syncytial virus hospitalization and wheezing/asthma decreases with age at follow-up (Régnier SA, et al. Pediatr Infect Dis J, 2013). To evaluate whether this also occurs between RV wheezing illness and the subsequent development of wheezing/asthma, we made an age based analysis to see whether there is any difference between less than 10 years old and more than or equal to10 years old. We have described this in the revised manuscript (Page 10, lines 8-23).

Comment 13: p10 line 24, meta-analysis does not create a larger sample size, in-fact yours (n=9) was smaller than each study.
Response 13: We thank you for this suggestion, we have changed our description to “Therefore, it is critical to systematically evaluate all relevant studies and to assess the overall association” in the revised manuscript (Page 10, lines 56-58; Page 11, lines 3).

Comment 14: Methods: "Besides" can be deleted p7 line9, this approach can be called jackknifing.
Response 14: We have deleted the word “Besides” in the revised manuscript (Page 6, lines 53).

Comment 15: Stop using the ">" sign for subgroup identification. This might help eliminate the error made in discussion.
Response 15: We have changed the ">" sign to “more than” in the revised manuscript.

Reviewer 4:
Comment 1: This is a reasonably well-written meta-analysis of the association between the rhinovirus wheezing illness and childhood asthma from 9 cohort studies. The results of the fixed-effect meta-analysis showed an increased odds of wheezing/asthma associated with rhinovirus illness in the first 3 years of life. The strengths of the study include its primacy, and use of cohort studies. However, there are a few specific comments and suggestions summarized below.
Response 1: We thank the reviewer for a thorough review of the manuscript and for the positive comments.

Comment 2: Did the authors register their review with PROSPERO http://www.crd.york.ac.uk/PROSPERO/ or another similar database prior to commencing their review? Similar to RCTs, a research protocol should be drafted and published for systematic reviews and meta-analyses to ensure scientific rigor.
Response 2: We have not registered this review with PROSPERO prior to commencing this analysis. However, we conducted this study according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We would like to register with PROSPERO or other similar database in advance when we perform a new meta-analysis in the future.

Comment 3: The search strategy is not detailed enough - for example in PubMed you should be using MeSH terms where appropriate and similar for the other databases. It is necessary to undertake a comprehensive search of the literature using an appropriate search strategy for each database searched. This ensures that the search is in fact systematic, reproducible and identifies all of the relevant literature. A complete search for at least one of the databases should be included as

supplementary material. I suggest speaking to a librarian or a similarly qualified individual. Further, a qualification of the search strategy should be provided (if undertaken), for example did the authors consult librarians?
Response 3: We thank the reviewer for this helpful suggestion. We have modified the search terms and established searching strategies with the help of a library expert of our school. An example of search details using Embase has been shown in supplementary table S1 of revised manuscript (Page 5, lines 26-31).

Comment 4: How did you deal with articles presenting results from the same study? I assume that this is what 'redundant information' pertains to in figure 1.
Response 4: We thank you for inquiring this issue. We have included 15 articles in this meta-analysis. Of them, 10 articles report the results of 4 longitudinal cohort studies with different follow-up periods. Only the study with longest follow up duration was included, this yielded 9 studies were included in overall analysis and subgroup analysis by ethnicity. In the analysis stratified based on age at follow-up, according to one reviewer's suggestion, we included all the studies with all follow up periods with total 15 studies. We have modified the results section in the revised manuscript (Page 9, lines 16-25).

Comment 5: Use of Newcastle-Ottawa Scale is controversial as the summary scores involve inherent weight of component items (Greenland 1994 "Quality scores are useless and potentially misleading" American Journal of Epidemiology, 300-301); the authors should discuss the limitations of these scores in their discussion
Response 5: We thank the reviewers for the suggestion. Use of NOS might be controversial as the summary scores involve inherent weight of component items, although it is a commonly used tool for quality assessment of non-randomized studies included in a systematic review and/or analysis. We have discussed the limitations of NOS in the discussion of the revised manuscript (Page 12, lines 38-46).

Comment 6: In the methods section the authors state that they assumed the HR and RR approximated the OR. This is only true if the outcome is rare. Is asthma considered to be a rare outcome (I wouldn't think so)? If not, this is not appropriate.
Response 6: We assumed that OR=HR=RR according to the study published previously (Greenland S, Epidemiol Rev, 1987). While this is only true if the outcome is rare, and asthma might not be considered as a rare outcome, so this assumption is inappropriate. By reviewing the relevant literature, we used RRs as the common measure of association across studies, directly considered HRs as RRs and transformed ORs into RRs using the formula RR=OR/ [(1-P0) (P0×OR)] (Zhang J, et al, JAMA, 1998). We have modified the methods sections in the revised manuscript (Page 6, lines 48-53).

Comment 7: The authors rely on the use of "statistical significance"; I strongly suggest that the authors read the recent paper by Greenland et al "Statistical tests, P values, confidence intervals and power: a guide to interpretations". European Journal of Epidemiology. 2016; 31:337-350 and that they remove any references to "statistical significance".
Response 7: We thank you for this helpful suggestion. We have read the paper by Greenland et al. and removed "statistical significance" in the revised manuscript.

Comment 8: The authors state that heterogeneity was confirmed from a p-value of 0.1 or I2 of more than 50%; how was this chosen?
Response 8: We calculated the heterogeneity among studies using Q statistic (significance level of $P < 0.10$) and I2 statistic (greater than 50% as evidence of significant inconsistency) according to the Cochrane Handbook for Systematic Reviewers of Interventions. We have cited this handbook as a reference in the revised manuscript (Page 7, lines 16; Reference 31).

Comment 9: The authors indicate that they carried out a sensitivity analysis by sequentially excluding each study; a reference is needed for this type of sensitivity analysis. This type of sensitivity analysis is not recommended; studies should be excluded for reasons chosen a priori.
Response 9: We thank the reviewer for this suggestion. We carried out a sensitivity analysis by sequentially excluding each study according to the study by Ferretti G et al (Ferretti G et al, Prog Lipid Res, 2015). By reviewing the relevant literature, we have modified the method of sensitivity analysis and excluded the studies in which OR was transformed into RR according to the study by Willi C et al (Willi C et al, JAMA, 2007) (Page 6, lines 53-58).

Comment 10: More information is needed in the Table 1. I suggest including the number of participants in each cohort, the number of participants with asthma, the follow-up time, the statistical method used, the confounding variables adjusted for in the analysis, and the amount of missing data. If necessary, some of this information could be provided in a supplementary table.
Response 10: We thank the reviewer for these suggestions. We have provided detailed informations of included studies in supplementary table S2 (Page 7, lines 51-53).

Comment 11: In the first paragraph of the discussion, the authors mention confounding. The main problem with cohort studies is confounding and I don't feel that this issue has been appropriately handled in this manuscript. A further discussion of potential confounding is needed, as is a deeper exploration of how each individual study has dealt with confounding in their analysis. Please refer to Egger et al "Spurious precision? Meta-analysis of observational studies". BMJ. 1998 pages 140-144 for further reading.
Response 11: We thank the reviewer for this helpful suggestion. We have read the paper by Egger et al and discussed the potential confounding factors in the revised manuscript (Page 10, lines 23-43).

Comment 12: In the discussion the authors mention that the diagnosis of wheezing/asthma was not always provided by a physician; I wonder if this would create a bias and whether it is worth carrying out a subgroup analysis of papers with/without physician diagnosis of wheezing.
Response 12: Some of studies included in our meta-analysis defined the outcome as physician-diagnosed wheezing/asthma. While in the studies reported by Ruotsalainen, Rubner, Lukkarinen and Teeratakulpisarn, wheezing/asthma diagnosis was made according to one of several criteria including diagnosis provided by physicians, this might affect the accuracy of the diagnosis and create a bias. In addition, these authors did not provide the precise number of participants diagnosed by physicians, so we are unable to carry out a subgroup analysis of papers with/without physician diagnosis of wheezing/asthma.

Comment 13: The authors state that there is no evidence of publication bias based on funnel plot and Egger's test; however Egger's test is known to have low power for meta-analysis of fewer than 20-25 studies. This meta-analysis includes only 9 studies; the limitations of this test should be discussed.
Response 13: We thank the reviewer for this helpful suggestion. Potential publication bias was assessed by funnel plot and Begg's and Egger's tests. We have provided the funnel plot as figure 5 and discussed the limitation of publication bias in the revised manuscript (Page 9, lines 36; Page 12, lines 46-51).

Comment 14: In light of my previous comments about confounding, the last line of the conclusion needs to be refined and made more specific.
Response 14: We thank the reviewer for this suggestion. We have revised the last sentence of conclusion to make it more specific (Page 13, lines 10-15).

Comment 15: In figure 1, it is not clear to me why 3 studies "with no subgroup analysis" were excluded. Can't these studies be included in the overall analysis?
Response 15: The major aim of the present study is to address the association between RV-induced

early wheezing and childhood wheezing/asthma. While the authors of the above 3 studies did not clearly indicated the identity of RV infection in individual participants in their studies, RV was only put into the group of non- respiratory syncytial virus or picornavirus, we could not get the number of RV infection, so we excluded these studies.

Comment 16: More details of the methods used to synthesize the data (i.e. the meta-analysis methods) should be included in the abstract.
Response 16: We thank the reviewer for the suggestion. We have added some details of the methods used to synthesize the data in the abstract of the revised manuscript (Page 2, lines 23-26).

Comment 17: It is not necessary to include the name of the software package in the abstract.
Response 17: We have deleted the information about software package in the abstract of the revised manuscript.

Comment 18: Reword the sentence on line 34 “There was a significant…” to “RV wheezing illness in the first 3 years of lie was associated with an increased odds of wheezing/asthma in later life (OR = 2.02; 95% CI = 1.58 - 2.57, P<0.001).”
Response 18: We thank the reviewer for this helpful suggestion. We have changed our description to “RV wheezing illness in the first 3 years of lie was associated with an increased odds of wheezing/asthma in later life” (Page 2, lines 33-36).

Comment 19: Reword the end of the sentence on line 43 from “…was only observed among < 10 years studies” to “was only observed for studies that had less than 10 years of follow-up”.
Response 19: We thank you for suggestion, it has been changed (Page 2, lines 40-41).

Comment 20: The last sentence of the conclusion “Larger-scale and well-designed studies are…” is too vague.
Response 20: We thank the reviewer for this suggestion. We have changed our description to “Large-scale and well-designed studies that adequately address concerns for potential confounding factors are required to validate the risk identified in the current meta-analysis” (Page 2, lines 48-51).

Comment 21: Page 4 Line 30 - RV-related diseases “are” globally…
Response 21: We are sorry for this issue. We have added “are” in this sentence (Page 4, lines 26).

Comment 22: Page 4 Line 38 - reference needed for sentence ending with “… for virus replication”.
Response 22: We have provided a relevant reference for this sentence in the revised manuscript (Page 4, lines 33; Reference 12).

Comment 23: Page 5 line 10 - reword “no study has been published which has combined systematic review and synthesis of the evidence.” to “no study has been published that has systematically reviewed the literature and synthesized the available evidence”.
Response 23: We thank the reviewer for this suggestion. We have changed “no study has been published which has combined systematic review and synthesis of the evidence.” to “no study has been published that has systematically reviewed the literature and synthesized the available evidence” in the revised manuscript (Page 4, lines 58; Page 5, lines 3).

Comment 24: Page 5 line 12/13-we performed “a systematic review and” meta-analysis…
Response 24: We have changed this sentence according to your suggestion (Page 5, lines 4).

Comment 25: Indicate the start date of the search no just the end date.
Response 25: We are sorry for these issues. We have added the start date of the search in the revised manuscript (Page 2, lines 21; Page 5, lines 24).

Comment 26: Page 6 line 31/32 - “outcome measure” should be changed to “effect estimate”.
Response 26: We thank the reviewer for this suggestion. We have changed “outcome measure” to “effect estimate” in the revised manuscript (Page 6, lines 28).

Comment 27: Page 6 line 52 -What do you mean by “by using data reported in the relevant literature”. Please clarify and include appropriate references.
Response 27: In the cases, where a study did not provide HR/RR/OR, we calculated the unadjusted RR using the formula RR =P1/P0, P1 indicated the incidence of the outcome of interest in the exposed group and P0 in the non-exposed group. We have changed our description and cited a relevant reference (Page 6, lines 44-48; Reference 27).

Comment 28: Page 7 line 9 - include categorizations for ethnicity and age at follow-up.
Response 28: We thank the reviewer for these suggestions. We have mentioned the categorizations for ethnicity (Caucasians and Asians) and age at follow-up (less than 10 years and more than or equal to 10 years) in the revised manuscript (Page 7, lines 18-21).

Comment 29: Table 1- present age in the same unit for every study (e.g. months).
Response 29: The age in every study has been presented by months in the table 1 of the revised manuscript (Page 8).

Comment 30: Table 1 is difficult to read. I suggest putting it in landscape format.
Response 30: We have revised the table 1 and put it in landscape format (Page 8).

Comment 31: Page 9 line 24/25-change the wording “… was only observed among < 10 years studies…” to “… was only observed among studies that had a follow-up time of less than 10 years…”. I would also suggest presenting the results for studies with follow up time > 10 years in this paragraph.
Response 31: We have changed our description and added the results of studies that the age at follow-up of participants was more than or equal to 10 years in the revised manuscript (Page 9, lines 21-25).

Comment 32: Page 9 line 39/40- include the funnel plot as a figure.
Response 32: We have provided the funnel plot as figure 5 in the revised manuscript (Page 9, lines 36).

Comment 33: Page 9 line 58 -change the wording “… was only observed among <10 years studies.” To “… was only observed among studies that had a follow-up time of less than 10 years”.
Response 33: We have changed our description according to our results and your suggestion (Page 10, lines 18-20).

Comment 34: Page 10 line 22 - sentence ending with “…development of subsequent wheezing and asthma.” needs a reference.
Response 34: We thank the reviewer for this suggestion. We have cited the relevant references in the revised manuscript (Page 10, lines 56; References 14 15 18-23 25 34-39).

Comment 35: It is helpful to show how many records were identified in each database and how many duplicate records were excluded.
Response 35: We thank the reviewer for this suggestion. We have provided more details of study selection for records in figure 1 of the revised manuscript (Page 7, lines 40).

Comment 36: Item #5 in the PRISMA checklist should be marked as 'Not reported
Response 36: We have marked the item #5 in the PRISMA checklist as "Not reported".

VERSION 2 – REVIEW

REVIEWER	Scott Montgomery Örebro University, Sweden
REVIEW RETURNED	21-Dec-2016

GENERAL COMMENTS	I am satisfied by the authors' responses to my queries. One further suggestion would be to alter the word 'race' on page 10 in the discussion to 'ethnic origin' - it will then be consistent with the details of the stratified analysis in the methods section.

REVIEWER	Ryan Pohlig University of Delaware, United States of America
REVIEW RETURNED	30-Dec-2016

GENERAL COMMENTS	Authors failed to address multiple concerns raised by this reviewer (#3). Example 1--Reaffirming what was done does not address the suggestion of performing a sub analysis using a common end point for studies when possible. Comment 2: Authors should consider running a second analysis using a common follow-up time and not just last one reported and include results if different or mention if same. Response 2: We thank the reviewer for this helpful suggestion. We have included all relevant studies (n=15) containing the same cohorts with different follow-up periods to perform subgroup analysis stratified by age at follow-up and provided the results in the revised manuscript (Page 9, lines 16-25). Example 2--Stating that all were categorical does not address criticism of not including continuous outcomes or weaknesses inherent in categorizing continuous outcomes. Comment 4: Importantly, authors should also considered non-categorical outcomes, and converted them to either an OR/HR/RR or converted all effect sizes into a standardized form (i.e. Hedge's). Response 4: We thank the reviewer for this suggestion. The outcomes in studies included in our meta-analysis were all categorical variables. Additionally, in addressing comment 10 ("We performed subgroup analysis stratified by race to evaluate whether different genetic backgrounds influenced the association between RV wheezing illness in the early life and subsequent development of wheezing/asthma."), their justification remains weak. Running separate analyses by race would not even answer address their hypothesizing that race moderates the relationship between wheezing & asthma. In order to test this an interaction effect of race by wheezing would need to be included in the model. I would reaffirm that since the OR/RR were essentially the same, that this sub-anlysis does not add anything to the paper. This reviewer also wants to know how statistical significant testing

	was performed in the sub analyses on groups of 2? This author only knows of 1 method for garnering CI estimates for n=1 and n=2 but you have to assume a distributional shape (reference: http://www.jstor.org/stable/2685995?seq=1#page_scan_tab_contents). It is quite possible that there is method for doing so unknown to this reviewer, but referencing another meta-analysis is not sufficient justification for performing statistical tests.

REVIEWER	Dr Amalia Karahalios Melbourne School of Population and Global Health The University of Melbourne, Australia
REVIEW RETURNED	02-Jan-2017

GENERAL COMMENTS	I have a few outstanding major comments about the manuscript: 1. The authors state that “10 articles reported the results of 4 longitudinal cohort studies with different follow-up periods.” The name of the cohort study should be included in Table 1. I have concerns about including multiple papers from the same study in a meta-analysis. The estimates from one study will be correlated, which will result in standard errors (and corresponding 95% confidence intervals) that do not reflect the true variability (i.e. they are too precise). If the authors choose to include multiple articles/papers from the same study, they need to account for the correlation in their analysis. However, I recommend that the authors do not pool estimates from the same study and instead include studies with the longest follow-up period. 2. The authors state that they derived an unadjusted relative risk when it was not presented in the paper. For studies that are at risk of confounding (e.g. cohort studies) it is not appropriate to pool unadjusted relative risks with adjusted relative risks. The authors need to go back to the original researchers and request adjusted relative risks. 3. The authors used the formula provided by Zhang et al 1998 to estimate the relative risk from the odds ratio. The authors should indicate what value of P0 was used in this equation. 4. The authors should discuss the amount of heterogeneity that was observed in each meta-analysis and explicitly state whether a fixed-effect or random-effects meta-analysis was used to pool the data. When a random-effects meta-analysis was used, which estimate of heterogeneity was used? How was this chosen? 5. I’m not convinced by the authors comment about the Newcastle Ottawa Sclae specifically, “components have so little weight that the variations might be submerged in the total score.” Minor comments – 1. Study design should be included in Table 1 or Table S2 2. Spell out any acronyms: a. 95% CI the first time it is used, which is in the data extraction in the methods section. b. lnRR should be spelled out – logarithm of the RR. c. nM should be spelled out Abstract: 1. The interpretation of the estimates is not quite right – for example, the authors state in the abstract “increased odds of wheezing/asthma in later life”, however they present a relative risk. This needs to be fixed. 2. Please replace “statistical association was identified in

	Caucasians" with "an increased risk of asthma wheezing/asthma was evident in Caucasian…" Methods 1. Include a reference to the PRISMA checklist in the methods section. 2. In the response to the reviewers' comments, the authors state that the search strategy was guided by a librarian – this should be stated in the manuscript as well. 3. The flowchart in figure 1 does not correspond to the search criteria detailed in the methods section. Specifically, the number of citations from the Chinese National Knowledge Infrastructure (CNKI), and Wanfang databases are not displayed. Please update figure 1. 4. Fixed-effects should be changed to fixed-effect Discussion 1. There are a number of grammatical errors in the discussion. 5. The authors state: "The main problem with cohort studies is confounding, which may distort results." Confounding might lead to biased results; distort is an odd description. Please change this.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:
Comment 1: One further suggestion would be to alter the word 'race' on page 10 in the discussion to 'ethnic origin' - it will then be consistent with the details of the stratified analysis in the methods section.
Response 1: We thank the reviewer for this suggestion. We have changed "race" to "ethnic origin" in the discussion of the revised manuscript (Page 12, lines 40).

Reviewer 3:
Comment 1: Reaffirming what was done does not address the suggestion of performing a sub analysis using a common end point for studies when possible. Authors should consider running a second analysis using a common follow-up time and not just last one reported and include results if different or mention if same.
Response 1: We thank the reviewer for this helpful suggestion. We have conducted a second analysis using a common follow-up time, and the results indicated that the association between early life RV wheezing illness and subsequent development of wheezing/asthma still remained significant (RR=2.00, 95% CI=1.71-2.33, P <0.001). We have added the results in the revised manuscript (Page 9, lines 15-18).

Comment 2: Stating that all were categorical does not address criticism of not including continuous outcomes or weaknesses inherent in categorizing continuous outcomes.
Response 2: We thank the reviewer for this helpful suggestion. In the studies included in our meta-analysis, disease outcome was often measured in non-categorical variables, such as the frequency of wheezing episodes and lung function. However, researchers always converted these non-categorical variables into a categorical variable (healthy condition, with childhood wheezing/asthma or not), which would cause the loss of information. We have discussed above limitation in the revised manuscript (Page 12, lines 20-28).

Comment 3: Additionally, in addressing comment 10 ("We performed subgroup analysis stratified by race to evaluate whether different genetic backgrounds influenced the association between RV wheezing illness in the early life and subsequent development of wheezing/asthma."), their justification remains weak. Running separate analyses by race would not even answer address their hypothesizing that race moderates the relationship between wheezing & asthma. In order to test this

an interaction effect of race by wheezing would need to be included in the model. I would reaffirm that since the OR/RR were essentially the same, that this sub-analysis does not add anything to the paper.
Response 3: We agree with the reviewer's point that running a subgroup analysis by race would add nothing to paper as having only 2 studies in Asians should not result in any model tests of significance. We have removed the results of subgroup analysis by race and discussed above limitation in the revised manuscript (Page 12, lines 35-43). We will carry out the subgroup analysis by race again to evaluate whether different genetic backgrounds influence the association between RV wheezing and wheezing/asthma risk when more relevant studies are available.

Comment 4: This reviewer also wants to know how statistical significant testing was performed in the sub analyses on groups of 2? This author only knows of 1 method for garnering CI estimates for n=1 and n=2 but you have to assume a distributional shape (reference: http://www.jstor.org/stable/ 2685995? seq=1#page_scan_ tab_ contents). It is quite possible that there is method for doing so unknown to this reviewer, but referencing another meta-analysis is not sufficient justification for performing statistical tests.
Response 4: We performed Fixed Mantel-Haenszel test in the sub analysis on a group of 2 which has been widely used in meta-analyses (Xiang Z et al. Oncotarget, 2016; Glazov G, et al. Acupunct Med, 2016). However, as you suggested, having only 2 studies in Asians should not result in any model tests of significance and referencing another meta-analysis is not sufficient for performing statistical tests, we have removed the results of subgroup analysis by race.

Reviewer 4:
Comment 1: The authors state that "10 articles reported the results of 4 longitudinal cohort studies with different follow-up periods." The name of the cohort study should be included in Table 1. I have concerns about including multiple papers from the same study in a meta-analysis. The estimates from one study will be correlated, which will result in standard errors (and corresponding 95% confidence intervals) that do not reflect the true variability (i.e. they are too precise). If the authors choose to include multiple articles/papers from the same study, they need to account for the correlation in their analysis. However, I recommend that the authors do not pool estimates from the same study and instead include studies with the longest follow-up period.
Response 1: We thank the reviewer for this helpful suggestion. We have provided the name of the cohort studies in the table 1 of the revised manuscript (Page 8). In our meta-analysis, we included 15 articles. Of them, 10 articles report the results of 4 longitudinal cohort studies with different follow-up periods. We have performed two overall analyses using the values from the longest follow-up (n=9) and values from all relevant studies (n=15) which contained the same cohorts with different follow-up periods, respectively. Both results showed that RV wheezing illness in the first 3 years of life was associated with increased risk of wheezing/asthma (Page 9, lines 6-18). To evaluate whether the association between RV wheezing and development of wheezing/asthma would change with age, we included all relevant studies (n=15) containing the same cohorts with different follow-up periods to perform subgroup analysis stratified by age at follow-up. While the estimates from one study will be correlated, which will result in standard errors (and corresponding 95% confidence intervals) that do not reflect the true variability. We have discussed above limitation in the revised manuscript (Page 13, lines 8-13).

Comment 2: The authors state that they derived an unadjusted relative risk when it was not presented in the paper. For studies that are at risk of confounding (e.g. cohort studies) it is not appropriate to pool unadjusted relative risks with adjusted relative risks. The authors need to go back to the original researchers and request adjusted relative risks.
Response 2: We thank the reviewer for the suggestion. We have sent the original researchers emails to request adjusted RRs, but have not received any reply yet. We guess they might not calculate the adjusted RRs. We could only use the unadjusted RRs and discussed the limitation of this in the

revised manuscript (Page 13, lines 3-8).

Comment 3: The authors used the formula provided by Zhang et al 1998 to estimate the relative risk from the odds ratio. The authors should indicate what value of P0 was used in this equation.
Response 3: We transformed the ORs into RRs using the formula RR=OR/ [(1-P0) (P0×OR)], in which P0 is the incidence of the outcome of interest in the non-exposed group (the incidence of wheezing/asthma in the children without RV-induced early wheezing). We have provided the P0 values in supplementary table S2 (Page 6, lines 56). There were 3 studies that did not provide P0 (Lemanske, et al. J Allergy Clin Immunol, 2005; Rubner, et al. J Allergy Clin Immunol, 2016; Kusel, et al. J Allergy Clin Immunol, 2007).We have sent the original researchers emails to request relevant data, but have not received any reply yet. We could only use the P0 of the study by Jackson as the P0 of the studies by Lemanske and Rubner for all of these studies using the same longitudinal cohort with different follow-up periods. We also used the P0 of the study by Jackson as the P0 of the study by Kusel, because the subjects of these 2 studies had similar characteristics. Due to the limitation of P0 which may affect the accuracy of the results, therefore, we performed another analysis excluding all the studies which use the P0 to convert OR to RR, and the results indicated that the association still remained significant, which were consistent with the results of previous overall analyses. The limitation and shortage of original data has been discussed in the revised manuscript (Page 13, lines 3-8).

Comment 4: The authors should discuss the amount of heterogeneity that was observed in each meta-analysis and explicitly state whether a fixed-effect or random-effects meta-analysis was used to pool the data. When a random-effects meta-analysis was used, which estimate of heterogeneity was used? How was this chosen?
Response 4: We thank the reviewer for this helpful suggestion. We have added the amount of heterogeneity observed in each meta-analysis and explicitly stated which effect model was used to pool the data (Page 9, lines 10-11; Page 9, lines 23-26). We calculated the heterogeneity among studies using Q statistic (significance level of $P < 0.10$) and I2 statistic (greater than 50% as evidence of significant inconsistency) according to the Cochrane Handbook for Systematic Reviewers of Interventions. A random effect model was adopted when heterogeneity between studies was significant ($I2 > 50\%$, $P < 0.10$). We have mentioned it in the revised manuscript (Page 7, lines 16-23).

Comment 5: I'm not convinced by the authors comment about the Newcastle Ottawa Sclae specifically, "components have so little weight that the variations might be submerged in the total score."
Response 5: NOS is a commonly used tool for quality assessment of non-randomized studies included in a systematic reviewer and/or analysis, however, use it may be controversial as the summary scores involve inherent weight of component items. No matter estimated effect varies with the quality score or not, the analyst can skip the quality-score analysis and go straight to the quality-component analysis to find out which components are responsible for the variation or avoid the risk of misleading conclusions, therefore, some researchers think that quality-score analysis may be superfluous (Greenland S. Am J Epidemiol, 1994; Reference 48). We have changed our description of the limitations of NOS in the revised manuscript (Page 12, lines 53-58; Page 13, lines 3).

Comment 6: Study design should be included in Table 1 or Table S2.
Response 6: The design of every study has been added in supplementary Table S2.

Comment 7: Spell out any acronyms: a.95% CI the first time it is used, which is in the data extraction in the methods section. b. lnRR should be spelled out – logarithm of the RR. c. nM should be spelled out.
Response 7: We are sorry for these issues. We have provided the full forms of these acronyms in the revised manuscript (Page 6, lines 30; Page 7, lines 6-8; Page 10, lines 43).

Comment 8: The interpretation of the estimates is not quite right- for example, the authors state in the abstract “increased odds of wheezing/asthma in later life”, however they present a relative risk. This needs to be fixed.
Response 8: We are sorry for this issue. We have changed the description to “increased risk of wheezing/asthma in later life” in the revised manuscript (Page 2, lines 35).

Comment 9: Please replace “statistical association was identified in Caucasians” with “an increased risk of asthma wheezing/asthma was evident in Caucasian…”
Response 9: We thank you for this suggestion. One of the reviewers pointed out that it made little sense to run a subgroup analysis by race because there were only 2 studies in Asians. We have removed the results of subgroup analysis by race and discussed the limitation of shortage of relevant studies in the revised manuscript (Page 12, lines 35-43). We will carry out the subgroup analysis by race again to evaluate whether different genetic backgrounds influence the association between RV wheezing and wheezing/asthma risk when more relevant studies are available.

Comment 10: Include a reference to the PRISMA checklist in the methods section.
Response 10: We have provided a relevant reference for the PRISMA checklist in the methods section of the revised manuscript (Page 5, lines 15; Reference 26).

Comment 11: In the response to the reviewers’ comments, the authors state that the search strategy was guided by a librarian-this should be stated in the manuscript as well.
Response 11: We thank the reviewer for this suggestion. We established search strategies with the help of a library expert of our school. We have added this in the methods section of the revised manuscript (Page 5, lines 30-32).

Comment 12: The flowchart in figure 1 does not correspond to the search criteria detailed in the methods section. Specifically, the number of citations from the Chinese National Knowledge Infrastructure (CNKI), and Wanfang databases are not displayed. Please update figure 1.
Response 12: We have revised the figure 1 and provided more details of study selection in CNKI and Wanfang databases.

Comment 13: Fixed-effects should be changed to fixed-effect.
Response 13: We are sorry for this issue. We have changed “fixed effects” to “fixed effect” in the revised manuscript (Page 7, lines 20).

Comment 14: There are a number of grammatical errors in the discussion.
Response 14: We are sorry for these issues. We have invited my colleague proficient in English to carefully check the manuscript and correct all spelling and grammar errors.

Comment 15: The authors state: “The main problem with cohort studies is confounding, which may distort results.” Confounding might lead to biased results; distort is an odd description. Please change this.
Response 15: We have changed our description to “The main problem with cohort studies is confounding, which might lead to biased results” in the revised manuscript (Page 10, lines 20).

VERSION 3 – REVIEW

REVIEWER	Emily Karahalios The University of Melbourne, Australia
REVIEW RETURNED	23-Feb-2017

GENERAL COMMENTS	The authors have addressed my major comments. Some very minor comments below: 1. In supplementary table 2 – indicate `Not reported' when the P0 was not reported. 2. Fixed-effect and Random-effects are the correct terminologies - please correct throughout the manuscript. 3. In Table 1, the cohort designation is difficult to read. Just list the name of the study next to each publication.

VERSION 3 – AUTHOR RESPONSE

Reviewer 4:
Comment 1: In supplementary table 2-indicate `Not reported' when the P0 was not reported.
Response 1: We thank the reviewer for this suggestion. We have indicated "No reported" when the P0 was not reported by the original researchers in the supplementary table 2.

Comment 2: Fixed-effect and Random-effects are the correct terminologies - please correct throughout the manuscript.
Response 2: We are sorry for these issues. We have carefully checked the manuscript and used the correct form of terminologies "fixed-effect and random-effects" in the revised manuscript (Page 2, lines 26; Page 7, lines 21).

Comment 3: In Table 1, the cohort designation is difficult to read. Just list the name of the study next to each publication.
Response 3: We thank the reviewer for this helpful suggestion. We have listed the name of the study next to each publication in the table 1 of the revised manuscript (Page 8).