

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)	Daycare attendance and respiratory tract infections: a prospective birth cohort study
AUTHORS	Schuez-Havupalo, Linnea; Toivonen, Laura; Karppinen, Sinikka; Kaljonen, Anne; Peltola, Ville

VERSION 1 - REVIEW

REVIEWER	J. Kevin Yin Australia's National Centre for Immunisation Research and Surveillance, The Children's Hospital at Westmead; Sydney School of Public Health, Faculty of Medicine, The University of Sydney
REVIEW RETURNED	26-Oct-2016

GENERAL COMMENTS	<p>Schuez-Havupalo and colleagues conducted and reported an interesting prospective cohort study evaluating the burden of RTIs in young children. Recognizing the inherent limitations of observational study, their findings would allow for greater understanding of the health and economic burden of RTIs in childcare attendees. However, I outlined below a number of issues for considerations.</p> <ol style="list-style-type: none">1. A key potential confounder, which has not been adjusted in the analyses, is the length of 'exposure' by study arm i.e. the number of hours the children spent in childcare during the study period. In the first paragraph of Results section, the authors mentioned "...the vast majority of children in daycare spent over 5 hours per day at daycare"; but then how many days per week, if such information was collected in the study? This is a critical factor in relation to the risk of developing RTIs among childcare attendees, given the childcare attendance was hypothesized as the source of infection.2. Were data on household income collected? This would be an interesting potential factor to explore in the analyses as well.3. If data mentioned in Points #1 and #2 are not available, the limitations section in Discussion should be revised accordingly.4. The standard reporting guideline for observational studies, STROBE, was not used.5. There are a number of places in the manuscripts where the wording or meaning was a bit unclear e.g. the first sentence of Paragraph 3 in Results. Further editorial editing would improve the overall readability.6. Should the Paragraph 6 (starting with "Due to strict requirements...") in Discussion be moved to Results?
-------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

REVIEWER	Adam Irwin Great Ormond Street Hospital for Children, UK
REVIEW RETURNED	23-Mar-2017

GENERAL COMMENTS	<p>General comments and recommendations: This is a well written manuscript providing evidence for a well-recognised problem of interest to paediatricians and primary care physicians and I enjoyed reading it. The results provide reassurance to clinicians and thereby to parents regarding the natural history of acute respiratory tract infections (RTI) in young children entering child care. I would recommend some changes to the analysis and reporting, but believe the manuscript to be of sufficient rigour and interest to be published.</p> <p>Major comments: The study uses a prospective data collection in a large birth cohort, which is a robust and time-consuming method. The STEPS study is a substantial dataset and this is an interesting and appropriate use of the data.</p> <p>The most significant limitation appears to be the number of children for whom data were missing, and I am not convinced that the best approach to this was to exclude these children. It seems unlikely that these data were missing completely at random (an assumption for the complete case analysis performed). For example, children were excluded if they discontinued child care (page 5, line 23), though I realise this is a relatively small number in the study. Families may discontinue childcare precisely because their child has experienced an increase in the number of febrile illnesses, and so failing to include these children may underestimate the effect of childcare on RTI. Conversely, it may be that those families whose children did experience a significant increase in the number of RTI following the start of childcare were more inclined to fill in the symptom diaries, while those who didn't were less inclined (because they were less worried). In summary I think missingness is likely related to the outcome (number of RTIs post childcare) so an attempt should be made to account for this. I would propose a process of data imputation or sensitivity analyses to test the robustness of the findings.</p> <p>Abstract: The abstract appears clear and concise.</p> <p>Methods: The description of the STEPS study is clear. It would perhaps be helpful to define the dates of follow up for this study (and not just of recruitment to STEPS). Was follow up finally completed when the last recruit to the current study completed 9 continuous months of childcare or some other measure?</p> <p>Page 5, line 19 describing the physician-led follow up is not clear. "During RTIs part of the cohort" doesn't make sense. I couldn't understand why there was a difference in the types of follow-up.</p> <p>The description of the daycare groups appeared clear though the "home care" group (used in the abstract) was given various names throughout the manuscript and this was a little confusing ("comparison group", "control group"). I think home care group best describes this group and it would be helpful to use this consistently.</p>
-------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

It is disappointing that a large birth cohort study investigating RTI in young children isn't able to report the impact of smoking in the home. It is well recognised that smoking has a significant impact on the rate of hospital admission in children, and that recent reductions in smoking significantly reduced child admissions with RTI (Been et al Eur Resp J 2015).

I found it difficult to understand the description of the repeated measures variance analysis, and the identification 'chronological thresholds of significance'. An alternative approach would be that of linear regression, with the possibility of fitting time as a piecewise variable to best fit the data.

Results:

Results from the first paragraph could be summarised in a table with overall characteristics of the cohort according to childcare group. This table should also include an overall summary of the outcome measures (sick days, antibiotics, parental absence) for each childcare group. I found it hard to find this information which was only really available in table 2. What was the overall mean sick days, antibiotic use, and parental days off work for the 3 groups? The authors focus on the observed increase following the start of daycare (which is interesting, and well illustrated), but the overall figures, which I suspect are very similar between the groups, should also be presented. This in itself would reinforce the message that any increase in symptoms associated with type of daycare is transient and the overall impact is limited.

Table 1 is confusing, and in addition listing p values is of no real value. It should be removed

Page 11, line 10 "when analysing..." is also confusing. I think this is saying that there is a difference between sick days and antibiotic use in the DCC group versus the "baseline" group (the home care group). It would be much more interesting to know what this difference was (how big) rather than whether or not it 'reached significance' (how confident you are statistically that it is in fact different).

Figure 4. It would be interesting to see how parental absence changed in the home care group (given the statement that it failed to return to baseline in both daycare groups because of the fact that both parents were more likely in work in the older children).

Discussion:

Page 13, line 48. I am not sure how useful it is to explain that there was variation between the months.

Page 14, line 40. For the reasons stated above, the exclusion of the missing data is likely to have introduced bias, and so 'careful exclusion' of missing data is not a very appropriate term.

Conclusion:

I think the authors could assert more strongly that though there is a transient increase in sick days and antibiotic use, the impact over the entire period of early life is minimal.

Minor comments:

In general, there is too much emphasis placed on p-values. Further

	<p>p-values <0.0001 are not plausible (they reflect too much confidence in the tails of the distribution) Page 6, line 50. I don't understand the term 'day-specific chronological follow-up', nor the statement 'not all above defined data related to the fixed age frame of 0-2 years'. This may relate to my earlier question about the time period of data collection for included children.</p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: J. Kevin Yin

Institution and Country: Australia's National Centre for Immunisation Research and Surveillance, The Children's Hospital at Westmead; Sydney School of Public Health, Faculty of Medicine, The University of Sydney

Competing Interests: None declared

Schuez-Havupalo and colleagues conducted and reported an interesting prospective cohort study evaluating the burden of RTIs in young children. Recognizing the inherent limitations of observational study, their findings would allow for greater understanding of the health and economic burden of RTIs in childcare attendees. However, I outlined below a number of issues for considerations.

Comment:

1. A key potential confounder, which has not been adjusted in the analyses, is the length of 'exposure' by study arm i.e. the number of hours the children spent in childcare during the study period. In the first paragraph of Results section, the authors mentioned "...the vast majority of children in daycare spent over 5 hours per day at daycare"; but then how many days per week, if such information was collected in the study? This is a critical factor in relation to the risk of developing RTIs among childcare attendees, given the childcare attendance was hypothesized as the source of infection.

Response:

The length of daycare exposure was assessed in all questionnaires as per day (not per week). In Finland both parents tend to attend work after daycare initiation, and it is therefore highly unusual for children to attend daycare on an irregular basis. Therefore, the reported high daily exposure to the daycare environment in the daycare group can be extrapolated to high weekly exposure.

Comment:

2. Were data on household income collected? This would be an interesting potential factor to explore in the analyses as well.

Response:

This information was added to table 2. Low household income was inversely related to RTIs, which is in accordance with the increased RTI rates in children of mothers with higher education levels. A mention of this association was added to the Discussion, p. 17.

Comment:

3. If data mentioned in Points #1 and #2 are not available, the limitations section in Discussion should be revised accordingly.

Response:

We added 2 sentences about daily vs. weekly/monthly exposure to daycare in our study to the limitations paragraph in Discussion, p. 16. Data on household income are available, please refer to earlier discourse.

Comment:

4. The standard reporting guideline for observational studies, STROBE, was not used.

Response:

We reassessed our manuscript according to STROBE where applicable. We have attached the completed STROBE checklist to our submission (Editorial request 1).

Comment:

5. There are a number of places in the manuscripts where the wording or meaning was a bit unclear e.g. the first sentence of Paragraph 3 in Results. Further editorial editing would improve the overall readability.

Response:

We have edited all possibly unclear parts of the manuscript, particularly the Methods and Results sections. The first sentence of paragraph 3 in Results was moved to Statistical analysis, because it describes the conduct of data analysis and should be clearer in that context.

Comment:

6. Should the Paragraph 6 (starting with "Due to strict requirements...") in Discussion be moved to Results?

Response:

This paragraph was modified. Please see our response to the first comment of reviewer 2.

Reviewer: 2

Reviewer Name: Adam Irwin

Institution and Country: Great Ormond Street Hospital for Children, UK

Competing Interests: None declared

General comments and recommendations:

This is a well written manuscript providing evidence for a well-recognised problem of interest to paediatricians and primary care physicians and I enjoyed reading it. The results provide reassurance to clinicians and thereby to parents regarding the natural history of acute respiratory tract infections (RTI) in young children entering child care. I would recommend some changes to the analysis and reporting, but believe the manuscript to be of sufficient rigour and interest to be published.

Comment:

Major comments:

The study uses a prospective data collection in a large birth cohort, which is a robust and time-consuming method. The STEPS study is a substantial dataset and this is an interesting and appropriate use of the data.

The most significant limitation appears to be the number of children for whom data were missing, and I am not convinced that the best approach to this was to exclude these children. It seems unlikely that these data were missing completely at random (an assumption for the complete case analysis performed). For example, children were excluded if they discontinued child care (page 5, line 23), though I realise this is a relatively small number in the study. Families may discontinue childcare precisely because their child has experienced an increase in the number of febrile illnesses, and so failing to include these children may underestimate the effect of childcare on RTI. Conversely, it may be that those families whose children did experience a significant increase in the number of RTI

following the start of childcare were more inclined to fill in the symptom diaries, while those who didn't were less inclined (because they were less worried). In summary I think missingness is likely related to the outcome (number of RTIs post childcare) so an attempt should be made to account for this. I would propose a process of data imputation or sensitivity analyses to test the robustness of the findings.

Response:

We agree that the major limitation of this study is missing data, which is probably not entirely random. However, for the following reasons we argue that factors unrelated to the effects of daycare on RTIs appear to be implicated with regard to missing information:

- Of the 676 children excluded, all were lacking daycare-related data (lacking a precise time of enrolment into daycare, or unknown home-care status at the defined age points, in addition to earlier drop-out or other lacking information in some cases), and 267 children were lacking information on RTIs. Thus, the limiting factors in terms of information were dates of enrolment into daycare and information on explicitly stated home care. Importantly, these data (e.g. date of enrolment) were mostly gathered at the time of start of daycare, indicating that most of the missing children left the study before the start of daycare. Our outcomes cannot be directly related to drop-outs that occurred before the follow-up of RTIs after daycare initiation. We have extended our evaluation on missing data in the text (Discussion, p. 16).

- data on daycare (from questionnaires and diaries) and data on RTIs (from diaries only) were extracted from different sources, which partially led to missing data. We assume that this cause of missing data should be unrelated to RTIs.

- The STEPS study is a sizeable prospective cohort study designed by a multiprofessional team of researchers (Lagström H et al. Int J Epidemiol 2013;42:1273-84). A large volume of data are obtained from a variety of different sources at numerous time points, and we therefore consider bias due to anticipation of our specific research question unlikely. It seems that a more likely reason for missing data is that the study was rather demanding for the families because of detailed questionnaires.

We carried out a review of our data, and unfortunately we were able to detect a reporting error with regard to the 27 children who finished daycare during follow-up. Contrary to our figure 1, these children were not excluded per se, but analyses excluded the follow-up after discontinuation of daycare. (Of the 27 children in question, 11 were in FDC and 13 in DCC, the remainder were excluded due to missing daycare-related information). We have performed an in-depth review of the processes giving rise to this error, and have no reason to doubt the integrity of the revised data. The error was corrected in figure 1. In conclusion, children who finished daycare during follow-up were not excluded, but follow-up time after discontinuation of daycare was. We performed sensitivity testing by additional variance analyses excluding all data from these children and the results remained consistent (added to Statistical analyses, p. 8, and to Results, p. 9).

Comment:

Abstract:

The abstract appears clear and concise.

Methods:

The description of the STEPS study is clear. It would perhaps be helpful to define the dates of follow up for this study (and not just of recruitment to STEPS). Was follow up finally completed when the last recruit to the current study completed 9 continuous months of childcare or some other measure?

Response:

We added to the first paragraph of Methods (Study population and conduct) the following sentence: "In this study, follow-up included an age frame of 0-2 years, until March 2012". We also modified Statistical analyses paragraph to clarify the follow-up.

Comment:

Page 5, line 19 describing the physician-led follow up is not clear. "During RTIs part of the cohort" doesn't make sense. I couldn't understand why there was a difference in the types of follow-up.

Response:

Part of the cohort were included in a study arm with more detailed infection follow-up, which comprised a standardized history, physical examination, and viral swabs during RTIs. Data obtained from the study clinic were not directly applied in this study, but they served as control information facilitating the overall evaluation of reliability with regard to RTI-data in the STEPS study. We added a reference (Toivonen L et al. Pediatrics 2016;138:e20161309) to that sentence.

Comment:

The description of the daycare groups appeared clear though the "home care" group (used in the abstract) was given various names throughout the manuscript and this was a little confusing ("comparison group", "control group"). I think home care group best describes this group and it would be helpful to use this consistently.

Response:

We have changed the wording to 'home care group' throughout the manuscript.

Comment:

It is disappointing that a large birth cohort study investigating RTI in young children isn't able to report the impact of smoking in the home. It is well recognised that smoking has a significant impact on the rate of hospital admission in children, and that recent reductions in smoking significantly reduced child admissions with RTI (Been et al Eur Resp J 2015).

Response:

Prevalence of maternal smoking in Finland is relatively low at 15% during early pregnancy, and it tends to be associated with lower socioeconomic status (Ekblad M et al. Eur J Public Health 2014;24:544-51). In our study, socioeconomic status was higher than in the general population, and we documented only a small number of smoking parents. We agree that it is disappointing that our study was not able to report some of the previously documented associations.

Comment:

I found it difficult to understand the description of the repeated measures variance analysis, and the identification 'chronological thresholds of significance'. An alternative approach would be that of linear regression, with the possibility of fitting time as a piecewise variable to best fit the data.

Response:

We have clarified the paragraph Statistical analysis in Methods. We suggest that repeated measures variance analysis in combination with the SAS MIXED procedure lends itself particularly well in this situation (exploring chronological aspects in a complex setting where every individual child commences daycare at a different time and parts of individual follow-up may have to be excluded selectively).

Comment:

Results:

Results from the first paragraph could be summarised in a table with overall characteristics of the cohort according to childcare group. This table should also include an overall summary of the outcome measures (sick days, antibiotics, parental absence) for each childcare group. I found it hard to find this information which was only really available in table 2. What was the overall mean sick

days, antibiotic use, and parental days off work for the 3 groups? The authors focus on the observed increase following the start of daycare (which is interesting, and well illustrated), but the overall figures, which I suspect are very similar between the groups, should also be presented. This in itself would reinforce the message that any increase in symptoms associated with type of daycare is transient and the overall impact is limited.

Response:

We have added the suggested table (new Table 1) and rewritten the related text in Results.

Comment:

Table 1 is confusing, and in addition listing p values is of no real value. It should be removed

Response:

We removed this table from the manuscript and have revised the text accordingly (page 11).

Comment:

Page 11, line 10 “when analysing...” is also confusing. I think this is saying that there is a difference between sick days and antibiotic use in the DCC group versus the “baseline” group (the home care group). It would be much more interesting to know what this difference was (how big) rather than whether or not it ‘reached significance’ (how confident you are statistically that it is in fact different).

Response:

Means and standard deviations of outcome measures are now given in table 1. The above-cited sentence was modified (p. 11).

Comment:

Figure 4. It would be interesting to see how parental absence changed in the home care group (given the statement that it failed to return to baseline in both daycare groups because of the fact that both parents were more likely in work in the older children).

Response:

Parental absences from work could not be compared to the home care group, since the stay-at-home parent was not employed in these families, and could therefore not be absent from work. This clarification was added to the second paragraph of Results (p. 9).

Comment:

Discussion:

Page 13, line 48. I am not sure how useful it is to explain that there was variation between the months.

Page 14, line 40. For the reasons stated above, the exclusion of the missing data is likely to have introduced bias, and so ‘careful exclusion’ of missing data is not a very appropriate term.

Response:

Sentences about variation between the months were removed from the Discussion. The part of Discussion about missing data has been changed (p. 16), as well as the last bullet point paragraph of the Article summary; Strength and limitations of this study, on p. 3.

Comment:

Conclusion:

I think the authors could assert more strongly that though there is a transient increase in sick days and antibiotic use, the impact over the entire period of early life is minimal.

Response:

The Conclusion paragraph (p. 17) has been changed according to the comment.

Comment:

Minor comments:

In general, there is too much emphasis placed on p-values. Further p-values <0.0001 are not plausible (they reflect too much confidence in the tails of the distribution)

Response:

The Results section was partly rewritten taking this comment into account. We now give the smallest p-values as $p < 0.001$.

Comment:

Page 6, line 50. I don't understand the term 'day-specific chronological follow-up', nor the statement 'not all above defined data related to the fixed age frame of 0-2 years'. This may relate to my earlier question about the time period of data collection for included children.

Response:

The paragraph Statistical analysis has been rewritten (p. 7-8).

VERSION 2 – REVIEW

REVIEWER	Adam Irwin Great Ormond Street Hospital for Children
REVIEW RETURNED	30-May-2017

GENERAL COMMENTS	<p>General comments: I think the authors have addressed my comments well and I have no significant remaining concerns. The flow diagram illustrates the 'missing' cases well and offers some reassurance that these excluded cases had only limited impact on the findings.</p> <p>Minor comment: Page 4, line 55 I think "lose" should be "loose" Page 6, line 10 "during RTIs part of the cohort" would be better understood as "during RTIs, part of the cohort..."</p>
-------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------