Author’s response to reviews

Title: The relationship between breastfeeding and reported respiratory and gastrointestinal infection rates in young children

Authors:

Nicole Frank (nicole.frank@ucdenver.edu)
Kristian Lynch (kristian.lynch@epi.usf.edu)
Ulla Uusitalo (Ulla.Uusitalo@epi.usf.edu)
Jimin Yang (jimin.yang@epi.usf.edu)
Maria Lönnrot (maria.lonnrot@uta.fi)
Suvi Virtanen (suvi.virtanen@thl.fi)
Heikki Hyöty (heikki.hyoty@uta.fi)
Jill Norris (jill.norris@ucdenver.edu)

Version: 1 Date: 26 Jun 2019

Author’s response to reviews:

Dear Editors of BMC Pediatrics,

We appreciate the reviewers’ comments and suggestions. Their comments helped us to clarify key points on how data were collected, on biases that may be present, on variables that we adjusted for in models, as well as interactions by gender, continent and season of visit. We trust that we have adequately addressed each of the questions and suggestions provided by the reviewers, and that the paper has been strengthened by the revision process.

Below you will find a point-by-point response to the comments from the reviewers. Changes in the manuscripts are indicated by track changes.

Best regards,

Nicole Frank
Reviewer reports:

Caroline Lodge (Reviewer 1):

Major comments

Methods

1. The timing of exposure and outcomes are not clear - would it be possible to further explain the temporal relationship of these variables. It is not clear if breastfeeding was categorized at the start of the 3 month period and infections after this - during the 3 month follow-up. These variables were obtained by the "TEDDY book questionnaire". Were discrete times recorded for cessation of exclusive and partial breastfeeding and for the occurrence of infections in each 3 month time interval? If so, why not use these times to explore the relationships between breastfeeding and infections?

Response:

Prior to every three month visit, the participating families were asked to record the infections since the last three month visit. To our knowledge, this is one of the few studies to have a diary book returned to a data collector at a visit on a regular basis for at least four years. This presented some unique challenges. For example, many of the extracted ICD-10 codes from our database were reported within a week of each other, or even on the same day. These were collapsed into a single infectious episode, as outlined in a previously published paper to which our reader is directed in the Methods section, and which is listed in our Reference section (Reference number 35). In addition, parents tended to report date of infections in the days leading up to a clinical visit and were less likely to report infections in the days just after the visit. To minimize the influence of short time ‘recall’ bias and any residual over-counting of infections, we opted to focus on whether they had a specific infection or not (yes, no) between visits, from the last visit and up to and including the day of the current visit, rather than focus on the exact timing within the three-month period.

Breastfeeding status was determined at the start of the 3 month interval during which infectious episodes were reported. In other words, infectious episodes were analyzed in 3-month intervals, and a participant was deemed to be “breast fed” during that interval if they were still breast fed at the start of that three month interval. This has been clarified in the Methods session, which now reads:

The odds of infections in a three-month period among children who were breastfed at the last visit (i.e. who were breastfeeding at the start of the three-month interval) compared to children
who were not breast fed at the last visit (or not breast fed at the start of the three-month interval) were calculated from coefficients of marginal logistic regression models.

2. Classification of infections is said to be from self-reported data - yet one of the outcomes is an ICD code (Line 216) Can you explain this further?

Response:

The parents reported their illnesses including fever in a TEDDY book diary. At each visit, the parents would bring their diary and the data collector would translate the reports into ICD-10 codes and enter them into the study database. The ICD10 codes were extracted from the database for the purpose of analysis. This was done to ensure better consistency of the self-reported infections across study sites in Europe and US.

The following has been added to the methods section as clarification:

The data collector reviews the reported illness symptoms, asks for clarifying information when appropriate, and assigns an ICD code or codes to each infectious episode, which are then entered into the study database. These codes are extracted from the database for data analysis.

3. The study has been going since 2004 and has reported on some 6 year follow-ups in other publications. Why does this investigation only span to a maximum of 48 months?

Response: For the first 4 years in the study (and of the child’s life), the diary book is submitted to the data collector at a clinical visits every three months. After the age of 48 months, visits occurred every three months only for children who were testing positive for islet autoimmunity (and who are therefore not included in this analysis). The remaining children switched to a schedule where clinic visits occurred only every six months after the age of 48 months. It is our intention to eventually examine long term follow-up up to 10 years and beyond. However because not all children have been followed that long as yet, and focusing on infections up to the age of 48 months allows for a consistent interval of data collection, we have chosen to only include data up to the age of 48 months in the analysis for this paper.

The following has been added to the Methods section to clarify the reasoning behind this decision:

Because the interval of reporting changes at the age of four years, for the purpose of this study, only follow-up data up to the age of four years will be included in the analysis.
4. Children not followed to 18 months were excluded - were there any differences for this portion of the cohort who were lost to follow-up?

Response: We have outlined in a previous publication that children who dropped out during the first year or two were more likely to be younger mothers with no father participation and female gender of the participating child. Also, mothers of children who withdrew were more likely to reduce their working hours, and were more likely to underestimate their child’s risk for T1D.

This previous publication has now been added to the References section. Additionally, the last paragraph of the discussion section has been re-written to better reflect the strengths and limitations of our study.

The following has been added to the methods section as clarification:

Nevertheless, despite the strengths of our study, there were limitations. Parental reporting of infections may still be prone to misclassification as only symptomatic infections were captured and there was difficulty differentiating between ICD10 codes of acute vs chronic infections. Also, there is likely to be some selection bias as the study excluded those participating families that were not as compliant with the TEDDY protocol. As we have previously reported, poor compliance or early loss to follow-up were related to a higher proportion of single, younger mothers, and mothers with fewer working hours during pregnancy (47). The influence of breastfeeding on parental reported infections may differ among this group of young families.

5. Children with antibodies were excluded due to perceived introduction of bias - did this subset have different associations for the exposures/outcomes being investigated?

Response: Few children develop islet autoantibodies before 9 months of age and so it would have little effect on associations with infections between the 3 to 6 months of age. We have not examined the effects of breastfeeding on infections after the child develops islet autoantibodies; however, we noticed the number of infections greatly increases once a mother is told their child has this biomarker and is at increased risk of T1D. As the goal was to focus on associations in a more general population, we decided not to examine this subgroup.

6. It is unclear why other variables were included in the models. Are all the variables included common cause confounders and do they have a plausible temporal relationship. For example - parental working status at 9 months is unlikely to be related to breastfeeding at 3 months. Season may be related to infection risk but is season related to breastfeeding?. Please indicate on what basis confounders were chosen to be added to the model. Was causal modelling theory involved, did the authors consider a directed acyclic graph?
Response: There were several reasons for including these factors. The first was to account for any possible effects due to missing or altered data that would create associations between breastfeeding and infections. While we agree with the reviewer that working at 9 months may not be associated with breastfeeding at 3 months, it is possible that working mothers spend less time directly observing their children, and are potentially less likely to report their data correctly, which would create association between these variables. Maternal working status may also impact a child’s transition to formula, if the working mother was not planning to continue exclusively breastfeeding or breastfeeding at all once they returned to work. Working status was evaluated at 9 months because that is the visit during which a socioeconomic questionnaire is completed.

Season at the end of the 3-month interval was included for a similar, but also for a more general, reason. Mothers in the summer, particularly in Europe, may have had more available time to answer questionnaires. However, alluding to what the reviewer was suggesting by considering a DAG, we were interested in adjusting for major risk factors of infections that may be unbalanced across breastfeeding groups within any particular between-visit period, regardless of whether season was, in theory, associated with breastfeeding and was a true confounder. We considered estimating the average causal effect of breastfeeding on infections by constructing a marginal structural model that would consider a propensity score on breastfeeding in any particular interval of time. However, we expect there will be very few additional observed factors that may be distributed unevenly across breastfeeding groups.

We believe that main caution with adjusting for other factors such as working status and season at the start of the interval was the possibility of introducing collider bias, i.e. infection and breastfeeding are both associated with the factor and thus adjusting for the factor would introduce a false association. The variables we chose are unlikely to be influenced by both breastfeeding and infections.

In the last paragraph of the discussion has been re-written to better describe the advantages and disadvantage of our study including the possibility of unobserved confounding and our reasoning for the factors for which we have adjusted.

7. Were any of the variables tested as modifiers - eg - sex, season. Most importantly, was centre tested. It is plausible that these relationships may differ by centre. I would like to see the relationships by centre before combining

Response: We have examined modifications by gender and continent of residence. The association of breastfeeding at 3 months with both febrile respiratory infection and ear infections, occurring between 3 and 6 months of age, was modified by gender, and this is now mentioned in the paper.
Are the associations between breastfeeding and reported infections modified by sex of child, place of residence or season?

Breastfeeding or length of breastfeeding and the association with infections after 6 months of age were not modified by gender, place of residence or season at last 3-month visit. The inverse association between breastfeeding at 3 months of age and respiratory infectious episodes with a reported fever (interaction, gender, \( p = 0.01 \); season age 3 months, \( p=0.02 \)), or with a reported otitis media (interaction, gender, \( p = 0.02 \), season at 3 months, \( p=0.02 \)) between 3 and 6 months of age, were both modified by sex of child and season when the child was 3 months of age. Among girls, breastfeeding was associated with a lower odds of both respiratory infectious episodes with a reported fever (yes breastfeeding vs no breastfeeding; \( \text{OR} \ 0.66, \ 95\%\text{CI} = 0.54 – 0.83 \)) and a reported otitis media (yes breastfeeding vs. no vs no; \( \text{OR} \ 0.55, \ 95\%\text{CI} = 0.41 – 0.74 \)).

No associations with these infections were seen among boys (respiratory infectious episodes with a reported fever; yes breastfeeding vs no breastfeeding; \( \text{OR} \ 0.98, \ 95\%\text{CI} = 0.80 – 1.21 \); a reported otitis media, yes breastfeeding vs. no; \( \text{OR} \ 1.02, \ 95\%\text{CI} = 0.77 – 1.35 \)). Similarly, when the child was 3 months of age, only if season was December to February was breastfeeding associated with a respiratory infectious episodes with a reported fever (yes breastfeeding vs no breastfeeding; \( \text{OR} \ 0.56, \ 95\%\text{CI} = 0.41 – 0.76 \) or with otitis media (yes breastfeeding vs no breastfeeding; \( \text{OR} \ 0.45, \ 95\%\text{CI} = 0.29 – 0.70 \)). At other 3-month seasons, no associations were observed (ORs >0.77). Site or continent of residence did not modify the associations.

8. I am a little confused by the statistical methods. The 3-6 month timepoint was investigated separately using a logistic regression model. All other timepoints with 3 month intervals from the 6 - 18 month timepoints were investigated using generalized estimating equations. Why were 2 different models used?. Why was the 3-6 month time period not included in the Gee with all the other times?

Does the Gee model assume that the relationship between breastfeeding and infection at each 3 month timepoint is the same? Is that a reasonable assumption?

Response: The GEE model was used to simply account for correlated data due to multiple visits from the same subjects. We have now clarified in the methods section that robust standard errors were used in the GEE and that we are using GEE simply to address the correlation issue. The following sentence now appears in the methods:

To account for the correlation of infections reported by the same family at multiple visits, the logistic models were estimated using Generalized Estimating Equations (GEE) with robust standard errors.
Results

1. The increase in respiratory infectious episodes reported is curious. Apart from the selective reporting suggested in the discussion, could this be related to reverse causation? Additionally, could this be related to the manifestation of infection. If all babies are equally exposed to respiratory infections, and breast fed babies are relatively protected, might you see an "increase" in uncomplicated urtis in these babies whilst less protected babies manifest these infections as laryngitis/tracheitis/conjunctivitis/Om etc?

Response:

We have incorporated these possibilities into our manuscript by adding the following sentences to the Discussion session:

Alternatively, mothers of children with more frequent respiratory symptoms may choose to breast feed for longer to impart to their children perceived health benefits derived from breast milk. It is also possible that respiratory illnesses in breastfeeding children tend to be less severe than those in children who do not breast feed, and therefore more frequently present as common cold, instead of manifesting as febrile illness, or otitis media.

2. It is not clear what has been adjusted for in table 2 and figure 2. Would the authors list all adjustments at the bottom of these tables/figures?

Response: This has been included in the footnote of Table 1 and Figure 2

All models were adjusted for gender, age of child, age of mother at birth, maternal education, single child, number of rooms in household, parental working and smoking status when child was nine months of age, country, if the child was a first degree relative of a type 1 diabetic individual, whether daycare or social group had started at the last visit, and season of the year at the start of the interval (i.e. last visit).

Discussion

It is claimed that exclusive breastfeeding has a greater protective effect than non-exclusive breastfeeding. Would the authors provide evidence that the odds ratios for these two groups are significantly different for the outcomes investigated?

Response: Thank you for pointing this out. While the effect size of exclusive breastfeeding and non-exclusive breastfeeding on odds of some infections was slightly different, it was not statistically so. Any mention of difference between exclusive and non-exclusive breastfeeding in
the 3-6 month time period has either been removed from the manuscript or modified to say it was similar.

Limitations

The select nature of the cohort has not been discussed in terms of generalizability of results

Line 361 - The authors describe the study size as beneficial for investigating interactions - it is not clear which interactions have been investigated in this study

Response: We have now removed this sentence related to benefit of interaction and have re-written a new paragraph to better reflect the strengths and limitations of the study and mention the interactions observed. We observed no interactions by site however we did see an interaction by gender and season of visit (see reply point 7 above).

The Conclusions state only the benefits for exclusive breastfeeding - these benefits are seen also relating to non-exclusive breastfeeding and the authors have not shown that these effects are significantly different. This statement needs to include both exclusive and non-exclusive breastfeeding.

Response: The reviewer is correct, and any mention of different associations based on type of breastfeeding (exclusive or non-exclusive) has been removed.

Minor Comments

3 decimal point in odds ratio in abstract. And line 258 - please change to 2

Response: This has been changed.

Numbers of children studied and outcome age to be included in the abstract

Response: This is now included.

The word "Conflictive" on the flow chart needs to be changed? did the authors mean conflicting?
Response: This has now been changed.

Karim Premji Manji, MBBS, MMED, MPH (Reviewer 2): The assessment of exclusive breastfeeding is not mentioned.

Response: Thank you for pointing this out. While the effect size of exclusive breastfeeding and non-exclusive breastfeeding on odds of some infections was slightly different, it was not statistically so. Any mention of difference between exclusive and non-exclusive breastfeeding in the 3-6 month time period has either been removed from the manuscript or modified to say it was similar.

Regarding the definition of exclusive breastfeeding for the purposes of this paper, this was already mentioned on page 7 (lines 145-146). However, further clarification has now also been added to the Methods section on page 9 (lines 190-191), which now reads:

Children who were a) exclusively breastfed (i.e. had not yet been introduced to formula or foods other than breast milk), b) breastfed but not exclusively (i.e. still received breast milk, but had also been introduced to formula and/or other foods) and c) no longer breastfed at 3 months of age were first examined in relation to respiratory and gastrointestinal infections between 3 and 6 month of age.

It would be a good idea to have the anthropometry assessed and reported as well.

Response: Although this would be interesting to investigate, this data has not been collected on our study participants and could not be analyzed in this study.