

26 June

Dear Dr. Smith

Manuscript ID BMJ.2017.039236 entitled "Impacts of London's road traffic air and noise pollution on birth weight: a retrospective population-based cohort study"

Thank you for sending us your paper, manuscript # XXX entitled "YYY" We sent it for external peer review and discussed it at our manuscript committee meeting. We recognise its potential importance and relevance to general medical readers, but I am afraid that we have not yet been able to reach a final decision on it because several important aspects of the work still need clarifying.

We hope very much that you will be willing and able to revise your paper as explained below in the report from the manuscript meeting, so that we will be in a better position to understand your study and decide whether the BMJ is the right journal for it. We are looking forward to reading the revised version and, we hope, reaching a decision.

Yours sincerely,

Rubin Minhas
Dr Rubin Minhas
BMJ Associate Editor
rm1000@live.com

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc.manuscriptcentral.com/bmj?URL_MASK=ccbe976f269e4e1680b2bc2826f01732

****Report from The BMJ's manuscript committee meeting****

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

Chair: Elisabeth Loder, Daoxin Yin, John Fletcher, Rubin Minhas, Julie Morris (Stats), Wim Weber, Tiago Villaneuva, Sophie Fletcher

Interesting RQ. 7-10 years old data, maybe difficult to get more recent data. Rescaled air pollution indices to interquartile range, becomes difficult to interpret effect sizes. Another source of confusion is role of confounding factors, e.g. smoking, no individual smoking data. The use of birth address - does this mean the address where the child was born, rather than the hospital? The effect size seems modest given the difference in birth weight versus use of interquartile range.

Assumption is most of time pre-delivery is present at home. The birth changes is dubious from a clinically meaningful viewpoint. Can they contextualise these changes or communicate the clinical significance better?

There are a number of specific issues relating to the statistical analysis and presentation of the data:

- Are the confounders included in the statistical model sufficient? Only COA level data were available for deprivation and tobacco information. Is the lack of individual smoking data a serious problem?

- The rescaling of air pollutant exposures to interquartile range increments makes it difficult for the general reader to interpret the effect sizes. Is this a conventional way of assessing air pollutant exposures?
- The actual effect sizes seem very small. For example, "IQR increases ...were significantly associated with mean term birth weight reductions of 7 to 13g" (Abstract). Taking as an example 'PM2.5 traffic exhaust' which has an IQR of 0.35 and mean 0.61 (Supplementary Table1), each IQR unit is 57% of the mean value. Hence, a greater than 50% change in the mean value is associated with birth weight reduction of 13g.
- There are a plethora of statistical significance tests. No acknowledgement of multiple comparisons is made.

Decision:

Put points - subject to addressing editors and reviewers comments.

Reviewer: 1

Recommendation:

Comments:

Summary

This study investigated the associations of air pollution and traffic noise with (low) birth weight and SGA in a more than 570,000 pregnant women living in Greater London and surrounding counties. The authors investigate potential confounding of associations with air pollution by traffic noise and vice versa.

The paper adds to the currently limited evidence on the association of traffic noise with pregnancy outcomes and the mutual confounding of air pollution and noise effects. It is well written and the large sample size and population-based study population are major strengths of the study.

Specific comments

Term births vs all birth: There seems to be some inconsistency. The abstract says that term (low) birth weight and SGA at term were the main outcome measures, whereas the method section states that associations were analyzed overall and in term births only without putting more emphasis pregnancy outcomes in term births. Figures and tables of associations of noise and air pollution with pregnancy outcomes in the main paper are provided for outcomes in term births only, which is consistent with the abstract, and associations in the overall population are summarized by a single sentence (page 12) without showing any of the results. I suggest to be consistent throughout the paper with regard to outcomes in term births being the main outcomes. Furthermore, associations for all birth should be presented in the supplemental material rather than stating "data not shown" as this increases transparency.

The paper would benefit from a map of the study area for readers that are not so familiar with the study area.

Air pollution exposures were available at 20m x 20m grid resolution, whereas noise exposures were assigned at address level. Potential exposure misclassification of especially air pollution from traffic, which is most consistently associated with birth weight in this study, may be a concern and should be discussed.

How was gestational age assessed? Based on ultrasound measures or based on last menstrual period? Please clarify.

Parity may be a potential confounder of the observed association. If information on parity is available, it should be included. If not, this should be discussed as a potential limitation.

How did you use tobacco expenditure in your models? Continuous? Categorical? Please clarify.
Statistical methods. Continuous vs categorical exposures: It is not clear why air pollution variables were used as continuous exposure variables and noise levels were categorized into 5 categories. Did you check the linearity of the exposure-response relationship and make the decision based on the outcome of these checks? If noise is linearly associated with the outcomes of interest, it is preferred to use it as a continuous variable, too, as categorization results in loss of power.

Since NO₂ and NO_x are perfectly correlated you may wish to consider leaving out one of the two.

Page 9, line 12. Add "singleton" after 572,910.

Page 11, line 16. "Protective" suggests causality and should be replaced by a more neutral expression e.g. "An increase in birth weight was observed with increasing O₃ exposure."

Results, page 12, line 36ff. Add that confidence intervals for trimester-specific effects largely overlap.

Results, description of Figure 3. Note that direction of noise effect changes upon adjustment for PM_{2.5} and PM₁₀.

Multicollinearity checks are mentioned in the methods section, but outcomes of these checks are not provided in the results section. Please add.

Results on exposure-ethnicity interaction should be shown (in the online supplement).

Spontaneous vs medical intervention deliveries were not adjusted for or used to exclude or stratify analyses. To the extent that medical intervention is related to clinical practice variation which could vary spatially because of hospital catchment area or cultural factors, this is a serious limitation that could potentially confound the spatial component of exposure metrics. Medical intervention could influence all of the outcomes, directly (pre-term, very preterm) or indirectly (term birth weight via gestation length and SGA if medical intervention influences the clinical estimate of gestational age). See for example: Gagnon et al. *Arch Gynecol Obstet.* 2013 Apr;287(4):633-9. Hanley et al. *Regional variation in the cesarean delivery and assisted vaginal delivery rates.* *Obstet Gynecol.* 2010 Jun;115(6):1201-8. Janssen et al. *Outcomes of planned hospital birth attended by midwives compared with physicians in British Columbia.* *Birth.* 2007 Jun;34(2):140-7. `

Table 1. Tobacco expenditure is missing here. Please add.

Table 2. This is probably minor given the high correlation between daytime and nighttime noise, but are air pollution effect estimates adjusted for daytime or nighttime noise? I suggest to have 2 separate tables one for daytime noise and one for nighttime noise, one of them could be moved to the supplement.

Add crude effect estimates to Figure 1.

Figures 1-3. Add a footnote with the exact increments for this analysis.

Supplementary Tables 2 and 3. Clarify increments for gestational age (is it 1 week?) and tobacco expenditure. Clarify that maternal age is in years.

Table 2 says "gestational age", table 3 says "gestational age (completed weeks)". What is correct? Revise accordingly.

Table 2, supplementary table 4-6. Replace "per IQR" by the exact increments used.

Additional Questions:

Please enter your name: Ulrike Gehring

Job Title: Associate Professor

Institution: Utrecht University, Institute for Risk Assessment Sciences

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

Reviewer: 2

Recommendation:

Comments:

The manuscript presents results from a large, registry-based study of traffic-related air pollution and traffic-related noise pollution and effects on several measures of birth weight (term birth weight, term low birth weight, term SGA). The study was conducted in London metropolitan area, using birth registry data for outcome evaluation and a dispersion modeling approach for exposure estimation.

The question of joint effects of noise and air pollution impacts of birth weight is an area that is not well studied, as the authors correctly identified, so this manuscript has the potential to add information to this topic. However, there are some key areas of the manuscripts that should be addressed. First, there is too much emphasis on statistical significance of the results. Instead, the focus should be on reporting on the estimated effects and associated confidence intervals. Second, it is not clear why the authors focused on these measures of birth weight. Biologically, is there evidence that a 7-13g reduction in birth weight is important to health (for birth weights across the distribution)? There is substantial evidence, however, that term LBW is an important predictor of infant (and future adult) health, so this would be my recommended outcome measure to use. In the Results section, the authors also mention doing an analysis for birth weight and SGA (not term), without explanation of the relevance of this analysis. Third, there are a number of studies that also examine the role of heat/temperature on birth outcomes, so that may be an important factor to consider.

Specific comments:

Methods - it is not clear why the study area was strictly limited to the area within the M25 boundary. Excluding the births in the areas overlapping the M25 would have excluded populations that were most highly exposed to pollution from this source. This would be an important group to include, if data are available.

Methods - Is the population-based metric of tobacco expenditures a good proxy for maternal smoking, or perhaps second-hand smoke exposure? Providing a reference that validates this as a proxy measure would be helpful.

Methods - p. 9 - For the births missing noise exposure and/or ethnicity (since these categories included 30,000+ records), it would be better to include them in the analysis and use imputation methods to address missing data.

Methods - p.9 - How were the cut points for noise determined? Are these biologically relevant categories?

Methods - p.9 - Were the adjustment factors in the models determined a priori, and why were these selected?

Discussion - p. 15 - Diesel PM includes a number of air toxics, so EC may not be the biologically active component causing the health outcomes of interest. Rather, EC is typically used as an indicator of diesel exhaust exposures, i.e. the exposure to the mixture of toxics contained in diesel exhaust.

General - The manuscript contains awkward phrasing in some areas (e.g. "Exposure is complex, close to roads and individual would be exposed to more primary exhaust and non-exhaust... particles...", "However, traffic also produces noise, which has been associated with e.g. hypertension and cardiovascular disease") and some grammatical issues that are distracting to the reader.

Additional Questions:

Please enter your name: Jo Kay Ghosh

Job Title: Health Effects Officer

Institution: South Coast Air Quality Management District

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

Reviewer: 3

Recommendation:

Comments:

This is a well-written manuscript, describing a well-conducted and comprehensive study to assess the relationship between near-road exposures to traffic-related air pollution and the potential for independent effects and/or confounding from traffic noise. This study adds to a growing and maturing literature suggesting causal relationships between air pollution and reduced term birthweight.

Air pollution exposures were estimated from a dispersion model. This model is likely more reliable for primary traffic vs related secondary emissions and perhaps also more accurate for gaseous vs particulate pollutants. Some of the authors have published land use regression models for multiple pollutants in

London and it would be helpful to evaluate if the observed associations are also evident with these other exposure estimates (which in other locations have proven to more accurately reflect measured concentrations than dispersion model estimates).

In addition there really needs to be some mention of model evaluation with measurements – the only reference is to the model website, not to any peer-reviewed publications, and no mention at all of evaluation. More importantly, what is the accuracy of the model estimates and are they equally accurate for the different pollutants that were evaluated? If not then this must be considered in the interpretation of the epidemiologic findings when comparing the different pollutant metrics. Treating all pollutant measures as equally accurate masks what could be important (and non-random) exposure misclassification.

What was the spatial resolution of the noise model? The manuscript suggest reduced exposure misclassification compared to other studies due to the use of addresses. This is valid but perhaps a bit oversold - the air pollution model has resolution of 20m so not fully resolved to each address and no indication is given on the resolution of the noise mode, except to indicate that estimates were generated for each address.

The air pollution metrics were analyzed as continuous exposures whereas the noise metrics are analyzed categorically. By treating the exposures differently it makes it rather difficult to accurately determine the impact of adjusting for one exposure vs the other – I'd suggest the authors pay more attention to this and would be well-advised to evaluate non-linear models for both exposures (including in joint exposure models) so one can address the shape of the relationships and evaluate whether the current categorization of noise exposures was appropriate. This issue could have bearing on the interpretation of the results.

The general conclusion of noise associations representing confounding is not entirely clear – the results (Figure 2) suggests attenuation of noise effects when air pollution included in models but there still does appear to be some suggestion of association for nighttime noise (especially in models including NO₂/NO_X). This result should get a bit more prominence in the Abstract and Discussion, or minimally some more discussion of why the authors conclude that all noise effects seem to be confounded. The logic as presented would seem to be: primary exhaust PM_{2.5} drives the air pollution associations (in the air pollution only models) so consequently the NO₂ associations in the joint (with noise) models are not given as much attention compared to the models with primary exhaust PM_{2.5}. This is not really valid either statistically or from a logic perspective – as it is more likely that primary PM_{2.5} and NO₂ reflect different aspects of the traffic pollutant mixture which may have different relationships with traffic noise.

As there were also associations present with non-exhaust air pollution it would seem that the manuscript should not be as forthright regarding the observed associations being related to diesel emissions and traffic exhaust (vs non-exhaust) findings (Abstract, What this Study Adds, Discussion).

The authors should tone down the mentions of this being the largest such study – large doesn't mean better and power is not really a concern for outcomes such as term birthweight or SGA. Both the Canadian and Danish studies, while smaller, were also population-based/included the entire population. It is OK to mention once that this is largest study but four times in the manuscript is a bit much.

There are multiple mentions in the manuscript of "significant" findings and also presents p-values. To conform to current best practices in epidemiology and with the statement of the American Statistical Association on statistical significance and p-values (<https://www.amstat.org/asa/files/pdfs/P-ValueStatement.pdf>) I'd suggest removing mention of significance and p-values throughout the manuscript. The authors have appropriately presented effect estimates and confidence intervals to allow for proper interpretation by readers.

There seems to be a missed opportunity in the study to separately examine the impact of aircraft noise measures such as those used in prior study conducted by several of the same authors (Hansell AL, Blangiardo M, Fortunato L, Floud S, de Hoogh K, Fecht D, Ghosh RE, Laszlo HE, Pearson C, Beale L, Beevers S, Gulliver J, Best N, Richardson S, Elliott P. Aircraft noise and cardiovascular disease near Heathrow airport in London: small area study. *BMJ*. 2013 Oct 8;347:f5432. doi: 10.1136/bmj.f5432.) Use of these measurements (and related modelled contours) would add much to the current analysis in

their ability to isolate noise from aircraft from the somewhat correlated traffic noise and traffic air pollution. The authors have included a sensitivity analysis where those with high exposures to aircraft noise were excluded (and not surprisingly found little impact on results given that this must represent a minority of the population) but much more useful would be a sensitivity analysis in which the subset of those exposed to high aircraft noise were specifically evaluated.

The authors may want to add to the Discussion the rather persuasive findings from a natural experiment which supports the suggestion of the importance of 3rd trimester exposures, as reported in the manuscript (Rich DQ, Liu K, Zhang J, Thurston SW, Stevens TP, Pan Y, Kane C, Weinberger B, Ohman-Strickland P, Woodruff TJ, Duan X, Assibey-Mensah V, Zhang J. Differences in Birth Weight Associated with the 2008 Beijing Olympics Air Pollution Reduction: Results from a Natural Experiment. Environ Health Perspect. 2015 Sep;123(9):880-7. doi: 10.1289/ehp.1408795)

Additional Questions:

Please enter your name: Michael Brauer

Job Title: Professor

Institution: The University of British Columbia

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here: I was senior author on a previous manuscript covering the same topic (in a different location).

Reviewer: 4

Recommendation:

Comments:

This is a well carried out study with good data linkage and coverage. This is a novel study, one of the largest to date on noise exposure and birth outcomes. The analyses are well executed, and a lot of data is presented. I feel that some of the interesting discussion has been constrained by the word-length of articles in the BMJ.

I would recommend publication of this paper in the BMJ after some minor changes to the paper.

Abstract

Consider revising the first few lines of the results section. It is very hard-going for someone coming anew to the paper, probably given the level of detail relating to the various metrics..which whilst I appreciate are important, I just don't think it reads very easily.

Conclusions: don't the findings that the relationship with noise apparently appears to reflect confounding by traffic air pollutant exposures conflict with the findings of recent reviews of combined exposures for other outcomes (Tétreault et al., 2013; Stansfeld, 2015). Why might this be? Is it because of something specific to the outcome here as the other papers tend to deal with cardiovascular outcomes? The reviews are limited in number of studies and size of studies. These issues should be elaborated on in the discussion.

Introduction

Could something brief about potential mechanisms for how air pollution and noise might biologically influence birthweight be added. I appreciate BMJ intros are short but this would benefit the reader for whom it may not be immediately obvious how noise influences birthweight especially.

Method

Section on noise exposures – I wasn't clear how moving house was taken into account in these estimates. This is later addressed in the discussion but should be clearly stated here. It is also worth checking the terminology in this section as I'm not sure that 'receptor placement' would be understood by a general audience.

Section on outcomes – I think these are the standard WHO definitions for birthweight but worth referencing. The last part of this section, where the N's are given for missing data – could % also be provided to help the reader understand more easily the extent of lost data.

Statistical methods – the categorisation of the Lnight – the lowest group is <50 dB. This is above the WHO NNG – was there not power to have a group that represented the NNG level of 40dB? I'm assuming not but could be worth discussing NNG in more detail.

Results

Are these levels of LBW and SGA what we would expect compared with national norms?

I wasn't very sure why the sensitivity analyses for those exposed to rail/air noise above 50dB included. Perhaps make the reasoning stronger here. The level of noise selected here at the lower end seems quite a low exposure and it seems a little crude to categorise everyone together who is 50dB or more. Similarly, whilst I appreciate that birthweight is associated with ethnicity, as noise and air pollution can be, I wasn't sure why we might expect the association between these exposures and birthweight to differ across ethnicities.

Discussion

I thought that noise got little discussion time in the discussion section. The key questions the paper raised for me that I didn't feel were well addressed were:

why would air pollution be more important for birthweight than noise?

why might air pollution only show an effect for birthweight, when for other outcomes such as cardiovascular health both exposures seem equally important and independent?

Is a -7 to -13g difference in birthweight important – I imagine it is but this isn't addressed.

The section where the results are compared to other findings is needed but it is quite long and not very easy to read for a lay reader. I think this section would benefit from some revision and cutting.

Limitations; the stronger findings for air pollution may be explained by differences between modelling quality for air pollution compared with noise. This should be discussed and listed as a potential limitation.

Additional Questions:

Please enter your name: Charlotte Clark

Job Title: Associate

Institution: Arup

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here: