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

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>SARS-CoV-2 cross-sectional seroprevalence study among public school staff in Metro Vancouver after the first Omicron wave in British Columbia, Canada</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Watts, Allison W.; Mâsse, Louise; Goldfarb, David; Irvine, Mike A.; Hutchison, Sarah M.; Muttucomaroe, Lauren; Poon, Bethany; Barakauskas, Vlėt E.; O’Reilly, Collette; Bosman, Else; Reicherz, Frederic; Coombs, Daniel; Pilblado, Mark; O’Brien, Sheila F; Lavoie, Pascal</td>
</tr>
</tbody>
</table>

VERSION 1 – REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Fredrik Methi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Norwegian Institute of Public Health</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>07-Feb-2023</td>
</tr>
</tbody>
</table>

| GENERAL COMMENTS   | 1. In this study the authors aim to compare the SARS-CoV-2 seroprevalence among school workers and a reference group of blood donors during the Omicron wave. Using cross-sectional seroprevalence sampling during the spring of 2022, they find no difference among school staff or the reference group. The outline of the study is sound, and their conclusions are supported by their data. The authors also nicely explain the basis of the study and how it was conducted. |
|                    | 2. My largest concern is that very little is written about the control group. In essence this article compares school staff with a control group of blood donors. However, throughout the article, the authors compare the school staff who answered the questionnaire with the school staff that were ultimately tested. More important is to establish whether the control group is similar to the school staff who got tested. The authors write that the control group is created using weights and provide a nice table in Supplemental Table 2. However, this is not sufficient. I would suggest the authors restructure the article so that Table 1 is the characteristics in both groups. In essence, I would combine Table 3 with Supplemental Table 2 to show that groups are similar, and hence build the study from there. |
|                    | 3. The authors should also be clearer on what is self-reported and not. If all characteristics are self-reported this should be included in the limitation section. For some things (e.g. mask wearing), people might tend to over report their own actions. |
|                    | 4. The authors write that school teachers are a highly vaccinated group: (1809/1830) had received at least two doses. If vaccination status in the control group (blood donors) is unknown, this could influence the findings. This should then be included in the limitations. In theory, 99% of school staff could be fully vaccinated and none of
the blood donors. If this was the case, then the similar rates of seropositivity would certainly indicate that school staff were more exposed and more at risk.

5. Also keep in mind that blood donors is an extremely healthy group of the population. This should also be written in the manuscript.

6. Finally, the authors claims that "Results are consistent with the premise that most infections were acquired outside the school setting, even with Omicron". However, authors also write that "antibodies persist in blood with assay sensitivity maintained until at least a year post-infection". This means that (if I understand correctly), in theory, all infections could be prior to the Omicron wave. And hence, the authors are not necessarily studying the Omicron strain, but rather how many persons were infected a year prior to date of testing - which happened to be during the Omicron wave. If this is correct, this should be addressed in the limitations section.

REVIEWER
Aleksander Galaś
Jagiellonian University, Department of Epidemiology and Preventive Medicine

REVIEW RETURNED
10-Feb-2023

GENERAL COMMENTS
Dear Editor,
Dear Authors,

Thank You for the opportunity to review statistical part of the manuscript entitled SARS-CoV-2 cross-sectional seroprevalence study among public school staff in Metro Vancouver after the first Omicron wave in British Columbia, Canada.
The content provided in the manuscript describes well the strategy applied by the authors, some issues however require clarification:
*there were models accounting for variations across individual schools which, as mentioned by the authors "produced comparable results". The school allocation should be used as a level in hierarchical design, unless it had no impact (questionable ...) therefore the authors should provide the data (which is currently "not shown") supporting the lack of variation across individual schools (general rule in evidence ... if data is not shown – don’t believe)
*the text 182 out of 205 positive controls ... and 10432 out of 10453 negative controls is mistaken. The number of negative controls was 10432 (21 were positive) and the total number of uninfected samples was 10453 (by https://diagnostics.roche.com/global/en/products/params/elecsys-anti-sars-cov-2.html). Probably the same issue is for 182 & 205. It should be corrected
*the key element in the Bayesian analysis is the level of informative priors which clearly impacts posterior. The authors should provide (in the appendix) the levels and arguments supporting their priors.
*regarding results obtained: the element which guarantees model accuracy is convergence. Visual inspection (as always) is very limited. Provide, please, convergence parameters (like Gelman-Rubin statistic).
*it is not clearly understandable why age and sex were omitted as covariates.
*the authors state “all analyses were done on cases with complete data”, so listwise procedure has been used. This usually leads to the decrease in the information and moreover, may be a source of bias, especially if the reason for missing data is not random (like it may
happened during pandemic). This should be discussed and explained in the manuscript.

And additionally, the authors state e.g. in the abstract “in the representative sample”. What supports the representativeness. There is no information about sampling (13497 eligible school staff, 2538 consented) and about participation rate.

**VERSION 1 – AUTHOR RESPONSE**

Reviewer: 1

1. In this study the authors aim to compare the SARS-CoV-2 seroprevalence among school workers and a reference group of blood donors during the Omicron wave. Using cross-sectional seroprevalence sampling during the spring of 2022, they find no difference among school staff or the reference group. The outline of the study is sound, and their conclusions are supported by their data. The authors also nicely explain the basis of the study and how it was conducted.

Thank you very much for these positive remarks.

2. My largest concern is that very little is written about the control group. In essence this article compares school staff with a control group of blood donors. However, throughout the article, the authors compare the school staff who answered the questionnaire with the school staff that were ultimately tested. More important is to establish whether the control group is similar to the school staff who got tested. The authors write that the control group is created using weights and provide a nice table in Supplemental Table 2. However, this is not sufficient. I would suggest the authors restructure the article so that Table 1 is the characteristics in both groups. In essence, I would combine Table 3 with Supplemental Table 2 to show that groups are similar, and hence build the study from there.

The blood donors data was a secondary data source that was not collected as part of our study but provided an excellent source data for comparative purpose. As a result we do have any control over the variables that were collected (this is the standard blood donor data collection procedures). We would also like to clarify that the blood donor seroprevalence was weighted to exactly match the school staff sample, for the residential neighbourhood postal code, age, sex and period of testing (by month). Therefore, a table comparing the school group with the blood donor group would be misleading and would not be representative of the blood donor data that was used to calculate the weighted seroprevalence in this group. In the supplemental file, we have clarified how the weighted analysis was conducted and we hope this will be clearer to the reader. Nonetheless, we agree with the reviewer that the school staff and blood donor groups could vary according to unspecified variables that could potentially influence seroprevalence estimates or exposure risk to the virus – we now acknowledge this as a potential limitation in the discussion. It is also important to reiterate also that the testing in both the school and blood donor groups were performed identically, using the same N-antibody assay, in the same lab, using the same methods for both samples (as our study samples for the school staff group was also carried out at the Canadian Blood Services lab). Thus, any differences due to testing should be miminized according to this testing strategy.

3. The authors should also be clearer on what is self-reported and not. If all characteristics are self-reported this should be included in the limitation section. For some things (e.g. mask wearing), people might tend to over report their own actions.

In the abstract and methods, we have tried to make it clearer which data comes from the online questionnaire and is thus based on self-report and this is now discussed as a potential source of bias in the limitations section.
4. The authors write that school teachers are a highly vaccinated group: (1809/1830) had received at least two doses. If vaccination status in the control group (blood donors) is unknown, this could influence the findings. This should then be included in the limitations. In theory, 99% of school staff could be fully vaccinated and none of the blood donors. If this was the case, then the similar rates of seropositivity would certainly indicate that school staff were more exposed and more at risk.

Remarkably, we know that the blood donor group was also 99% vaccinated based on self-reported data and this was also confirmed by Spike serology testing (not shown). Thus, both groups were highly vaccinated, so we do not think this issue is relevant. We have added this information to the method section of the article, and hope it addresses the reviewer’s point.

5. Also keep in mind that blood donors is an extremely healthy group of the population. This should also be written in the manuscript.

This is included in the discussion, but we have further highlighted this point in the limitation section. Blood donors are healthier than the general population, but this is likely also true of education workers. Essentially, a healthier control group would likely underestimate community seroprevalence estimates, and if true this would reinforce our conclusion?

6. Finally, the authors claims that “Results are consistent with the premise that most infections were acquired outside the school setting, even with Omicron”. However, authors also write that “antibodies persist in blood with assay sensitivity maintained until at least a year post-infection”. This means that (if I understand correctly), in theory, all infections could be prior to the Omicron wave. And hence, the authors are not necessarily studying the Omicron strain, but rather how many persons were infected a year prior to date of testing - which happened to be during the Omicron wave. If this is correct, this should be addressed in the limitations section.

Theoretically this may be possible. However, the reviewer could look at the following study: https://www.cmaj.ca/content/194/47/E1599 where seroprevalence estimates were obtained longitudinally in Vancouver, BC across multiple phases of the pandemic. In this study the authors show that infection-related seropositivity rates were about 10% in Sept-Oct 2021, before the Omicron wave, increasing to 42% by March 2022. This would suggest that most infections occurred after the emergence of Omicron. Also, though we clarify the temporal relationship of the recruitment in relation to the Omicron wave, we are not stating that the study seroprevalence estimate increase is due to Omicron specifically, agree?

Reviewer: 2

1. there were models accounting for variations across individual schools which, as mentioned by the authors “produced comparable results”. The school allocation should be used as a level in hierarchical design, unless it had no impact (questionable …) therefore the authors should provide the data (which is currently “not shown”) supporting the lack of variation across individual schools (general rule in evidence … if data is not shown – don’t believe)

We included the marginal posterior estimates of prevalence by school location in the supplemental under supplemental table 4, and in the main results, in the revised article.

2. the text 182 out of 205 positive controls … and 10432 out of 10453 negative controls is mistaken. The number of negative controls was 10432 (21 were positive) and the total number of uninfected samples was 10453 (by https://diagnostics.roche.com/global/en/products/params/elecsys-anti-sars-cov-2.html). Probably the same issue is for 182 & 205. It should be corrected
Overall the number of samples used to estimate specificity was 10453. These were taken from individuals before December 2019 so must all be treated as negative controls. Out of 10453, 21 were reactive so 10453 – 21 = 10432 tested negative out of 10453 who were negative. This may have come from a misinterpretation in how we applied these values, where they are used in the Bayesian model as an observation of a Binomial process with the sensitivity/specificity as the rate parameter and the total sample as the denominator.

3. the key element in the Bayesian analysis is the level of informative priors which clearly impacts posterior. The authors should provide (in the appendix) the levels and arguments supporting their priors.

The priors incorporated into the model are included in the appendix. The arguments for the priors to specify the variance of the sensitivity and specificity are included. Additional arguments for other priors are now also included.

4. regarding results obtained: the element which guarantees model accuracy is convergence. Visual inspection (as always) is very limited. Provide, please, convergence parameters (like Gelman-Rubin statistic).

The Gelman-Rubin statistic has now been reported in the results section.

5. it is not clearly understandable why age and sex were omitted as covariates.

The mixed-effects models now include age and sex as covariates; results did not change, but values have been updated and this is now described in the methods.

6. the authors state “all analyses were done on cases with complete data”, so listwise procedure has been used. This usually leads to the decrease in the information and moreover, may be a source of bias, especially if the reason for missing data is not random (like it may happened during pandemic). This should be discussed and explained in the manuscript.

We now include the amount of missing data in the methods section and clarify that there was almost no missing data (<0.8%) in the main outcome variables of interest here. Since it is so low, we have not included this as a limitation. Descriptively, the amount of missing data is displayed in the tables. We hope this answer is satisfying to the reviewer?

7. And additionally, the authors state e.g. in the abstract “in the representative sample”. What supports the representativeness. There is no information about sampling (13497 eligible school staff, 2538 consented) and about participation rate.

We compared the demographic information of the eligible staff with the participating staff in Supplemental Table 1. Differences among those who filled out the survey and completed the serology are now presented in Tables 1-3 so that the reader can see any differences. Our interpretation of the tables is that there are no notable differences, except for an increase in females in the Vancouver school sample compared to the corresponding population (we have added this clarification in the result section).

**VERSION 2 – REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Fredrik Methi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norwegian Institute of Public Health</td>
<td></td>
</tr>
</tbody>
</table>

| REVIEW RETURNED | 21-Mar-2023 |
**GENERAL COMMENTS**

Only minor comments:

Abstract line 16, testing results should be written in two words.
Page 4, line 31: Should it be “compared [to] the community seroprevalence”?
Page 6, line 38: “Seperate” should be written as “separate”
Page 12, line 9: “metropoliean” should be written as “metropolitan”
Reference 21 should be cited correctly.

**REVIEWER**

Aleksander Gałaś
Jagiellonian University, Department of Epidemiology and Preventive Medicine

**REVIEW RETURNED**

31-Mar-2023

**GENERAL COMMENTS**

Dear Editor,
Dear Authors,
The manuscript entitled “SARS-CoV-2 cross-sectional seroprevalence study among public school staff in Metro Vancouver after the first Omicron wave in British Columbia, Canada” has been improved after the implemented corrections.

Majority of the remarks provided in the first review have been addressed properly by the authors. In my opinion, however, there are still some issues which should be considered before publication:

- *strengths and limitations: non-random participation (as stated in 3rd point) violates the representativeness, therefore the second bullet point is not correct and should be removed. Remember, please, that “A representative sample is a group or set chosen from a larger statistical population or group of factors or instances that adequately replicates the larger group according to whatever characteristic or quality is under study.” Due to the study protocol used, and differences between Staff sample vs. Population (for instance in sex distribution) observed it is not proper to name the study sample as “representative school staff sample”.

- *introduction: statement “there has been very little-no published data” … slightly convincing, very little or no? verify and correct accordingly
- *authors have mentioned they excluded staff ‘on-leave’. Thus it means there is some possibility people on sick leave due to COVID-19 were excluded? Clarify, please.
- *negative controls are these who were controls and tested negative, so the proper construction of this sentence should be 182 positive out of total 205 infected samples and 10432 negative out of total 10453 uninfected samples. Surely, after the explanation the message is understandable but the explanation was not included into the manuscript.

**VERSION 2 – AUTHOR RESPONSE**

Specific response to issues raised by Reviewer#2:

Regarding the comment about the representativity of the sample, we have removed the term ‘representative’ from the Strengths and Limitations of This Study bullet points as suggested.

On the point about exclusion of those who were on-leave, we would like to clarify that the participants who were excluded were those who were on leave at the initial eligibility screening in early 2021. The numbers were small (28 out of 3,790 people screened were omitted because they reported they were
on leave = 0.7%). We did not collect the reason that participants were on-leave, but reasons could include mental health, pregnancy, child or elder care, sabbatical, extended travel, to obtain additional education/training, sickness, etc. Given that the seroprevalence at the time was low (2.3%) it seems extremely unlikely that these exclusions would have much impact on our findings now. However, those who reported being on-leave at the time of this analysis were included in the sample (n=29) as long as they enrolled in this current study phase in early 2022. We have reworded the corresponding section (first two paragraphs of method section, on page 4 of the manuscript) to make this clearer.