

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)	SARS-CoV-2 seroprevalence in students and teachers – a longitudinal study from May to October 2020 in German secondary schools
AUTHORS	Kirsten, Carolin; Unrath, Manja; Lück, Christian; Dalpke, Alexander H.; Berner, Reinhard; Armann, Jakob

VERSION 1 – REVIEW

REVIEWER	Kriemler, Susi University of Zurich, Epidemiology, Biostatistics and Prevention
REVIEW RETURNED	17-Mar-2021

GENERAL COMMENTS	<p>This study assessed seroprevalence in grade 8-11 students and teachers, including 13 secondary schools and 1538 students and 507 teachers in May 20 shortly after reopening of schools after the first wave of the pandemic, and in Sept/Oct 21 before the second wave of the pandemic. Seroprevalence was 0.6% and 0.2% in adolescents and teachers, respectively, defined as double confirmed antibody tests. 5/12 students had a history with a confirmed SARS-CoV-2 infection. There were neither any history of household transmission, nor clustering of seropositive students at school.</p> <p>This study is important and relevant as there are just a few school-based studies about seroprevalence in youth focussing on the school setting. And this is even more true for longitudinal data.</p> <p>The main limitations of the study are the focus on a convenient sample leading to a lack of generalisability, the lack of a power calculation to determine sample size needed, and the very scarce description of the epidemiological situation of Saxony that allows to set the context of the study.</p> <p>It is not clear how students, teachers and schools were selected from the general population of schools. Was it a convenience sample or were schools selected randomly? More information is needed to know about generalizability of findings: How many schools were eligible, how many contacted, how many participated. Where there differences between participating and non-participating schools. How was the participation rate of students within the total number of students per school and classes during both assessments etc. This is important to judge the precision of clustering within classes and schools. The only information we get is the between 21 and 573 students participated per school.</p> <p>The power calculation would enable the authors to be certain that the sample size was sufficient to answer the study questions.</p>
-------------------------	---

As the community setting in which the study takes place is very important, we need more details to understand whether the seropositivity was linked to the community setting or rather the school setting. The authors can reason about, as the first assessment did place right after the reopening of the school, thus seroprevalence was probably more a mirror of the home setting. The second assessment took place in sept/oct but we do not know about the evolution of the pandemic during the time before and also not whether there were holidays at schools in between (that prevented clustering), and whether they shut down in between etc. This second measure, if schools remained open, would indeed suggest that transmission in schools were extremely low.

In the introduction you write that numbers of children based on individual cases are lower, but on the population level, seroprevalence is very similar to adults – see Spain, UK, Switzerland, which simply reflects the pauci-symptomatic disease pattern in children and adolescents. Please adapt your statement.

You define seropositivity based on 2 positive serological test results. It may be interesting to see the numbers when only 1 test is used. I wonder whether you may have some quantitative MFI values and look at discordance and concordance based on the strength of antibody response?

An important information in your study is missing, namely how many adolescents were at home due to sickness or quarantine and might contribute to why the numbers were so low, especially during the September/October measurement. It is obviously possible that you were testing too early before the rise of cases during the second wave.

The discussion about the diagnosis of seropositivity is interesting. Yet, it may be overconservative to use a double confirmation of 2 tests as diagnostic criterium. It may also make sense to document results based on a single positive test criterium and look at the evolution. It might be too cautious in light of the variability, complexity and uncertainty of cut-offs for each serotest.

Different serological tests were used (e.g. Diasorin LIAISON etc): are there references? How is the sensitivity/specificity of used tests for the students of interest, or if not existing at least for adults? When calculating seroprevalence one should adjust for the hierarchical structure of the sample (schools, classes) and for sensitivity/specificity of the tests.

You write that clustering within schools was very low, but clustering is not defined. It may be interesting to reflect about the number of positive students per class as this might indeed be the place where transmission – if ever – takes place. Could you inform the reader whether there were several students seropositive on the class level?

The low transmission within families is indeed surprising. You may add some more contact tracing studies that clearly show transmission within families, and discuss reasons why you come up with different results?

T-cells may indeed play an important role in remaining or cross-reactive immunity. Can you explain a bit better to which of your

	<p>results you refer here?</p> <p>Table 1 describes the study sample and results. Were these the same students and teachers participating in round 1 and 2? Or were there new students and teachers participating in round 2 and drop outs after the first testing round?</p>
--	---

REVIEWER	Kadkhoda, Kamran Cleveland Clinic
REVIEW RETURNED	21-Mar-2021

GENERAL COMMENTS	<p>Great work but needs tweaking as follows:</p> <ol style="list-style-type: none"> 1. given the low specificity of these assays compared with PRNT, it important to either confirm their current results with a PRNT OR require BOTH Abbott and Euroimmun tests be positive to call an initial-positive/eq result as positive. Here's one great example: https://wwwnc.cdc.gov/eid/article/27/2/20-4088_article 2. Authors need to discuss their findings in the context of herd immunity as well. 3. Authors need to highlight that sero-epidemiological studies may not only suffer from low specificity but also from sero-reversion over time, therefore not really useful tools in public health decision-making as the notion of using COVID serology for sero-surveillance being a great choice, has rather turned into a cliché, in need of revamping.
-------------------------	---

REVIEWER	Lahner, Edith Universita degli Studi di Roma La Sapienza Dipartimento di Scienze Medico-Chirurgiche e Medicina Translazionale
REVIEW RETURNED	26-Mar-2021

GENERAL COMMENTS	<p>I read with interest this well-written paper on SARS-CoV2 seroprevalence in students and teachers showing a low infection rate.</p> <p>I have a few comments:</p> <p>I missed a paragraph on limits of the paper.</p> <p>It would be interesting to provide clinical information on Covid19 on dropped-out students and teachers. If this is not possible, I would suggest to discuss the lack of information on dropped-out study subjects.</p>
-------------------------	---

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

1) It is not clear how students, teachers and schools were selected from the general population of schools. Was it a convenience sample or were schools selected randomly? More information is needed to know about generalizability of findings: How many schools were eligible, how many contacted, how many participated. Where there differences between participating and non-participating schools. How was the participation rate of students within the total number of students per school and classes during both assessments etc. This is important to judge the precision of clustering within classes and schools. The only information we get is the between 21 and 573 students participated per school.

The schools were selected by the state office for schools and education ('Landesamt für Schule und Bildung (LASUB)' of the federal state of Saxony without involvement of the study team. All eligible students (grade 8-11) and all teachers at the respective schools were invited to participate via written information material handed out by the headmasters. After contacting the study team interested students and teachers were then enrolled in this study. Participation rates varied from 12%-50% per school.

In accordance with German data protection and privacy policies no contact was ever made between the study team and non-participating students and teachers. Therefore, we cannot make assumptions on differences between participating and non-participating schools or students/teachers. We added a paragraph to the method section providing context.

"Schools were chosen by the state office for schools and education without involvement of the study team and all eligible students and teachers were invited to participate at each school. Participation rates varied from 12%-50% per school."

2) *The power calculation would enable the authors to be certain that the sample size was sufficient to answer the study questions..*

A power calculation was not performed since there is no null hypothesis that could be rejected. This was set up as an epidemiological study aimed to provide longitudinal seroprevalence data. We performed a sample size calculation based on an expected seroprevalence of 1% with 5% precision and a 95% confidence level which yielded a minimum sample size of 500 participants which we exceeded at both timepoints.

3) *As the community setting in which the study takes place is very important, we need more details to understand whether the seropositivity was linked to the community setting or rather the school setting. The authors can reason about, as the first assessment did place right after the reopening of the school, thus seroprevalence was probably more a mirror of the home setting. The second assessment took place in sept/oct but we do not know about the evolution of the pandemic during the time before and also not whether there were holidays at schools in between (that prevented clustering), and whether they shut down in between etc. This second measure, if schools remained open, would indeed suggest that transmission in schools were extremely low.*

We added a paragraph explaining that there was no lockdown between the two visits.

"Between the two study visits schools in Saxony remained open with the regular summer break from July 20th until August 28th 2020."

4) *In the introduction you write that numbers of children based on individual cases are lower, but on the population level, seroprevalence is very similar to adults – see Spain, UK, Switzerland, which simply reflects the pauci-symptomatic disease pattern in children and adolescents. Please adapt your statement?*

We added the following sentence referring to the recent review of Lai et al. to our introduction.

"and a recent review on population-based seroprevalence studies found no evidence of overrepresentation of schoolchildren"

5) *You define seropositivity based on 2 positive serological test results. It may be interesting to see the numbers when only 1 test is used. I wonder whether you may have some quantitative MFI values and look at discordance and concordance based on the strength of antibody response*

We are happy to provide results based on a more liberal seropositivity definitions – as reviewer 1 prefers – and a more conservative definition – as reviewer 2 prefers. Given these differences in opinion and the fact that our approach matches the expected PPV in our population based on the test characteristics provided by the manufacturer we feel confident that our definition provides the most accurate results. We therefore would prefer to show these additional results in a supplemental table and not within the main manuscript since they do not change our main message – persistent low seroprevalence despite open schools.

While we have quantitative MFI values we do not feel confident to draw conclusions from this small sample size (12-25) and therefore would prefer not to present this data in an epidemiological study.

We added the following sentence in the results section;

“Using more liberal (≥ 1 test positive) or more conservative (3 tests positive) definitions for seropositivity does not change the persistent low seroprevalence in the study population.”

and we will provide the following supplemental table

Supplemental table 1: seroprevalence based on different seropositivity definitions

	first study visit (May/June)	second study visit (September/October)
≥ 1 serological test positive	22/2045 (1.1%)	25/1779 (1.4%)
≥ 2 serological tests positive	12/2045 (0.6%)	12/1779 (0.7%)
3 serological tests positive	9/2045 (0.4%)	5/1779 (0.3%)

6) An important information in your study is missing, namely how many adolescents were at home due to sickness or quarantine and might contribute to why the numbers were so low, especially during the September/October measurement. It is obviously possible that you were testing too early before the rise of cases during the second wave.

Unfortunately, we do not have accurate information on this issue since dropouts were not required to provide reasons for their discontinuation of the study. However, since both study visits took place in a low prevalence setting (7-day incidence rates ranged from 1 to 30/100.000 between May and October 2020) we are quite confident that quarantine measures were likely not responsible for the dropout rate.

We added the following paragraph in the results section.

“During the study period laboratory-confirmed SARS-CoV-2 infections per 100,000 inhabitants in Saxony increased from 139 to 245 and; 7-day incidence rates ranged from 1/100.000 to 30/100.000.”

7) The discussion about the diagnosis of seropositivity is interesting. Yet, it may be overconservative to use a double confirmation of 2 tests as diagnostic criterium. It may also make sense to document results based on a single positive test criterium and look at the evolution. It might be too cautious in light of the variability, complexity and uncertainty of cut-offs for each serotest?

Please see question 5.

8) *Different serological test were used (e.g. Diasorin LIAISON etc): are there references? How is the sensitivity/specificity of used tests for the students of interest, or if not existing at least for adults? When calculating seroprevalence one should adjust for the hierarchical structure of the sample (schools, classes) and for sensitivity/specificity of the tests*

Specificities and Sensitivities based on manufacturers labeling are now included in the method section. Performance of the primary test used (Diasorin) is discussed in the last paragraph of the discussion.

“In our population, a positive predictive value of 42.9% could be observed which was nearby an expected PPV of 45.3% for a prevalence of 0.59% population and the given test characteristics (sensitivity 97.6%, specificity 99.3%)”

9) *You write that clustering within schools was very low, but clustering is not defined. It may be interesting to reflect about the number of positive students per class as this might indeed be the place where transmission – if ever – takes place. Could you inform the reader whether there were several students seropositive on the class level*

Clusters in the German rules set up by the RKI are defined as at least 2 epidemiological linked cases. The initial 12 cases happened during the strict lockdown in March/April 2020. After the reopening of schools in May until the second study visit we could not detect any additional cases and therefor no clusters as well. We clarified that we sampled the same students and teachers in May/June and September/October in the method section.

“A second visit and repeat blood sampling **of the same participants** took place between September 15th and October 13th 2020.”

10) *The low transmission within families is indeed surprising. You may add some more contact tracing studies that clearly show transmission within families, and discuss reasons why you come up with different results*

We expanded our discussion adding more transmission studies as references.

“While close contact with COVID-19 patients—especially in the same household—has been shown to increase viral transmission [19], a review of household transmission studies found secondary attack rates of only 0.17 [20] with underage household members being less likely affected compared to adults. Our finding that only one out of 24 participants with a confirmed SARS-CoV-2 infection in the same household became indeed infected as measured by antibody production supports these findings as well as findings that children in general appear to be less susceptible to SARS-CoV-2 compared to adults [21, 22]. In addition, these results support studies showing that certain quarantine and separation measures than can effectively reduce the probability of viral transmission even in close contact situations [23].”

11) *T-cells may indeed play an important role in remaining or cross-reactive immunity. Can you explain a bit better to which of your results you refer here?*

We clarified this point in our discussion.

“The fact that we could not detect one additional seropositive participant in over 4 months is surprising even in a low prevalence setting, given that the reported cases doubled in the same period of time in Saxony. One explanation might be the recently reported detection of SARS-CoV2 spike-reactive CD4+ T cells could be detected in 35% of SARS-CoV2 unexposed healthy blood donors arguing for a certain level of T-cell crossreactivity. Such reactions could arise from exposure to commonly encountered Corona viruses. With children

being frequently exposed to common Corona viruses it might be hypothesized that they are less susceptible to SARS-CoV2 infection due to a background of T-cell crossreactivity [24].

12) Table 1 describes the study sample and results. Were these the same students and teachers participating in round 1 and 2? Or were there new students and teachers participating in round 2 and drop outs after the first testing round

We sampled the same students and teachers 2x – minus the dropouts. We clarified this in the method section.

“A second visit and repeat blood sampling of the same participants took place between September 15th and October 13th, 2020.”

Reviewer 2:

1) given the low specificity of these assays compared with PRNT, it important to either confirm their current results with a PRNT OR require BOTH Abbott and Euroimmun tests be positive to call an initial-positive/eq result as positive. Here's one great example:
https://wwwnc.cdc.gov/eid/article/27/2/20-4088_article.

Unfortunately we do not have PRNT data. We are happy to provide seroprevalence data based on a more conservative seropositivity definition requiring all 3 tests to be positive. However, given that reviewer 1 feels that requiring 2 positive tests is overconservative and the fact that our approach matches the expected PPV in our population based on the test characteristics provided by the manufacturer we feel confident that our definition provides the most accurate results. We therefore would prefer to show these additional results in a supplemental table and not within the main manuscript since they do not change our main message – persistent low seroprevalence despite open schools.

We added the following sentence in the results section.

“Using more liberal (≥ 1 test positive) or more conservative (3 tests positive) definitions for seropositivity does not change the persistent low seroprevalence in the study population.”

And we will provide the following supplemental table

Supplemental table 1: seroprevalence based on different seropositivity definitions

	first study visit (May/June)	second study visit (September/October)
≥ 1 serological test positive	22/2045 (1.1%)	25/1779 (1.4%)
≥ 2 serological tests positive	12/2045 (0.6%)	12/1779 (0.7%)
3 serological tests positive	9/2045 (0.4%)	5/1779 (0.3%)

2) Authors need to discuss their findings in the context of herd immunity as well.

We added the following sentence to our discussion:

“Herd immunity in the population of students and teachers appears not to contribute substantially to protection in a low prevalence setting.”

3) Authors need to highlight that sero-epidemiological studies may not only suffer from low specificity but also from sero-reversion over time, therefore not really useful tools in public health decision-making as the notion of using COVID serology for sero-surveillance being a great choice, has rather turned into a cliché, in need of revamping

While sero-reversion might impact longitudinal studies, we cannot provide evidence or context for this problem since there was indeed no sero-reversion during our study period. We samples the same population twice and found the same seropositive individuals.

Reviewer 3:

I missed a paragraph on limits of the paper.

It would be interesting to provide clinical information on Covid19 on dropped-out students and teachers. If this is not possible, I would suggest to discuss the lack of information on dropped-out study subjects.

Unfortunately, we do not have this information since dropouts were not required to provide reasons for their discontinuation of the study. We added a paragraph on limitations including the lack of information on dropped out study subjects:

“There are several limitations to our study. We cannot provide information on eligible but nonparticipating students and teachers in the selected schools requiring additional caution when generalizing these results. In addition, there is a relevant loss of participants in the follow-up sampling. While we do not have information why certain individuals dropped out, the fact that the second study visit took place in a before the beginning of the second wave (7-day incidence rates around 30/100.000) makes it unlikely that personal illness or widespread quarantine measures were responsible for this drop in participation.”

VERSION 2 – REVIEW

REVIEWER	Kriemler, Susi University of Zurich, Epidemiology, Biostatistics and Prevention
REVIEW RETURNED	24-Apr-2021

GENERAL COMMENTS	Reviewer 1: 1) It is not clear how students, teachers and schools were selected from the general population of schools. Was it a convenience sample or were schools selected randomly? More information is needed to know about generalizability of findings: How many schools were eligible, how many contacted, how many participated. Where there differences between participating and nonparticipating
-------------------------	--

	<p>schools. How was the participation rate of students within the total number of students per school and classes during both assessments etc. This is important to judge the precision of clustering within classes and schools. The only information we get is the between 21 and 573 students participated per school.</p> <p>The school were selected by the state office for schools and education ('Landesamt für Schule und Bildung (LASUB)' of the federal state of Saxony without involvement of the study team. All eligible students (grade 8-11) and all teachers at the respective schools were invited to participate via written information material handed out by the headmasters. After contacting the study team interested students and teachers were then enrolled in this study. Participation rates varied from 12%-50% per school.</p> <p>We still feel that you can get this information how schools were selected, how many secondary schools there are in total in the state of Saxony, and how the state office contacted the schools (all, a selection etc)</p> <p>2) The power calculation would enable the authors to be certain that the sample size was sufficient to answer the study questions..</p> <p>A power calculation was not performed since there is no null hypothesis that could be rejected. This was set up as an epidemiological study aimed to provide longitudinal seroprevalence data. We performed a sample size calculation based on a expected seroprevalence of 1% with 5% precision and a 95% confidence level which yielded a minimum sample size of 500 participants which we exceeded at both timepoints.</p> <p>Please integrate into the text. May be I'm wrong but this IS a power calculation.</p> <p>5) You define seropositivity based on 2 positive serological test results. It may be interesting to see the numbers when only 1 test is used. I wonder whether you may have some quantitative MFI values and look at discordance and concordance based on the strength of antibody response</p> <p>We are happy to provide results based on a more liberal seropositivity definitions – as reviewer 1 prefers – and a more conservative definition – as reviewer 2 prefers. Given these differences in opinion and the fact that our approach matches the expected PPV in our population based on the test</p>
--	---

	<p>characteristics provided by the manufacturer we feel confident that our definition provides the most accurate results. We therefore would prefer to show these additional results in a supplemental table and not within the main manuscript since they do not change our main message – persistent low seroprevalence despite open schools. While we have quantitative MFI values we do not feel confident to draw conclusions from this small sample size (12-25) and therefore would prefer not to present this data in an epidemiological study. We added the following sentence in the results section; “Using more liberal (≥ 1 test positive) or more conservative (3 tests positive) definitions for seropositivity does not change the persistent low seroprevalence in the study population.” and we will provide the following supplemental table</p> <p>Could you still confirm that the positive children in round 1 and 2 were the same? Or were there some that lost antibodies and newly acquired them?</p> <p>Minor: Legends in Tables are missing</p>
--	--

REVIEWER	Kadkhoda, Kamran Cleveland Clinic
REVIEW RETURNED	19-Apr-2021
GENERAL COMMENTS	Neutralization assays are key for these sero-surveys.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:

1) It is not clear how students, teachers and schools were selected from the general population of schools. Was it a convenience sample or were schools selected randomly? More information is needed to know about generalizability of findings: How many schools were eligible, how many contacted, how many participated. Where there differences between participating and nonparticipating schools. How was the participation rate of students within the total number of students per school and classes during both assessments etc. This is important to judge the precision of clustering within classes and schools. The only information we get is the between 21 and 573 students participated per school.

The school were selected by the state office for schools and education ('Landesamt für Schule und Bildung (LASUB)' of the federal state of Saxony without involvement of the study team. All eligible students (grade 8-11) and all teachers at the respective schools were invited to participate via written information material handed out by the headmasters. After contacting the study team interested students and teachers were then enrolled in this study.

Participation rates varied from 12%-50% per school.

We still feel that you can get this information how schools were selected, how many secondary schools there are in total in the state of Saxony, and how the state office contacted the schools (all, a selection etc)

There are 537 secondary schools in Saxony (schoolyear 2020/2021). The state office for schools and education ('Landesamt für Schule und Bildung (LASUB)' of the federal state of Saxony selected 13 of them and contacted them asking to participate in this study. None of the schools declined to participate. We added this information to the method section.

"After the reopening of the schools in Saxony on May 18th, 2020 students grade 8–11 and their teachers in 13 secondary schools in eastern Saxony were invited to participate in the SchoolCoviDD19 study. Schools were chosen by the state office for schools and education without involvement of the study team out of the 537 secondary schools in Saxony. Only the selected schools were contacted, none of them declined participation. All eligible students and teachers were invited to participate at each school. Participation rates varied from 12%-50% per school."

2) The power calculation would enable the authors to be certain that the sample size was sufficient to answer the study questions..

A power calculation was not performed since there is no null hypothesis that could be rejected. This was set up as an epidemiological study aimed to provide longitudinal seroprevalence data. We performed a sample size calculation based on a expected seroprevalence of 1% with 5% precision and a 95% confidence level which yielded a minimum sample size of 500 participants which we exceeded at both timepoints.

Please integrate into the text. May be I'm wrong but this IS a power calculation.

We integrated the paragraph into the method section.

"A sample size calculation was performed based on a expected seroprevalence of 1% with 5% precision and a 95% confidence level which yielded a minimum sample size of 500 participants which we exceeded at both timepoints."

3) You define seropositivity based on 2 positive serological test results. It may be interesting to see the numbers when only 1 test is used. I wonder whether you may have some quantitative MFI values and look at discordance and concordance based on the strength of antibody response

We are happy to provide results based on a more liberal seropositivity definitions – as reviewer 1 prefers – and a more conservative definition – as reviewer 2 prefers. Given these differences in opinion and the fact that our approach matches the expected PPV in our population based on the test characteristics provided by the manufacturer we feel confident that our definition provides the most accurate results. We therefore would prefer to show these additional results in a supplemental table and not within the main manuscript since they do not change our main message – persistent low seroprevalence despite open schools.

While we have quantitative MFI values we do not feel confident to draw conclusions from this small sample size (12-25) and therefore would prefer not to present this data in an epidemiological study.

We added the following sentence in the results section;

"Using more liberal (≥ 1 test positive) or more conservative (3 tests positive) definitions for seropositivity does not change the persistent low seroprevalence in the study population."

and we will provide the following supplemental table

Could you still confirm that the positive children in round 1 and 2 were the same? Or were there some that lost antibodies and newly acquired them?

11 initially seropositive participants continued to be seropositive at the second timepoint. 1 initially seropositive participant had only one positive serological test at the second timepoint and 1 participant with only 1 positive test result initially was positive in 2 serological test the second time. We clarified this in our results.

“one participant who tested positive in 2 assays in May tested positive in only one assay in October and was therefore no longer considered seropositive per study definition, while and one participant who had with equivocal results initially did test positive in two serological tests 3 months later. The remaining 11 seropositive participants had no changes in their test results.

4) Legends in Tables are missing

Legends were added to the tables.

Reviewer 2:

1) Neutralization assays are key for these sero-surveys

We agree that neutralization assays provide valuable information regarding potential immunity against SARS-CoV-2. However, we did not aim to make any assumptions in this regard. We just wanted to provide information on to what extent infections did take place in the study population during the course of the study. We don't want to present estimates on immunity or protection in our population, but on the rate of infections which could be confirmed by means of serology with 3 different assays.

VERSION 3 – REVIEW

REVIEWER	Kriemler, Susi University of Zurich, Epidemiology, Biostatistics and Prevention
REVIEW RETURNED	27-May-2021
GENERAL COMMENTS	Good job - accept as it is.