16 February 2021

Dear Dr Loder,

Thank you for your positive email. We are pleased to be able to address the comments raised by the editorial committee and the referees. We found the comments insightful and addressing them has improved both the substance and the presentation of the work. We include the detailed, point-to-point response to the editorial comments and both referees.

We also enclose the revised manuscript and supporting information that reflects these changes. We look forward to proceeding with this version.

On behalf of the coauthors,

Dr Robert Challen and Leon Danon

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

Present: Wim Weber (chair); Rafael Perera (statistician); Tiago Villanueva; David Ludwig; Joseph Ross; John Fletcher; Jessica Kimpton; Nazrul Islam; Clara Munro; Elizabeth Loder

* The methods section would benefit from additional information. The selection process of how you arrived at the 109545 matched patients is not very clear and requires additional explanation. (Fig. 1 is very confusing).

Response:

We have revised figure 1 and expanded the methods section - the details of which are given below.

* You find an average of 171 deaths out of 54773 patients in the S-gene negative arm of the study compared to 101 out of 54773 in the S-gene positive control arm. With the limited adjustments, we wonder how reliable this figure is. Might you comment? For example, it appears that you were not able to take comorbidities into account in your analysis.

Response:

The relatively small number of deaths included in the analysis is due to the fact that there is only a limited period of time when both S-positive and S-negative variants were present in the population, as the S-negative variant rapidly became dominant. We believe that a strength of this analysis is that we capture this transitionary period and thereby minimise for many of the confounders that might otherwise affect the results.

It is true that the data do not contain information on comorbidities, so we were not able to account for them directly in the analysis, other than by matching by geographical area, ethnicity, age and IMD, which capture some of the risk factors associated with comorbidities. With current understanding, there is no evidence or mechanistic reason why individuals with certain comorbidities should be more likely to be infected with one variant over the other. If anything, it seems plausible that individuals with comorbidities would be more likely to be in the data with the S-positive variant, which would incline the results towards the null.

We have this paragraph in the discussion:

"Some of the increased hazard could be explained by comorbidities. We were not able to directly adjust for comorbidities in this analysis. There is no information about comorbid conditions in the data we analysed, although this will be partly controlled by age, ethnicity and index of multiple deprivation. There is currently no evidence of a mechanistic reason why people with certain comorbidities would be infected with one variant or the other. However, it is possible that people with certain comorbidities are both at a higher risk of infection with VOC-202012/1 and have a higher mortality rate. This would tend to reduce the hazard ratio attributable to VOC-202012/1 alone."

* We discussed how to best put results in context. Although the HR of 1.7 seems alarming, the absolute risks remain low and increased transmissibility might be more worrisome than an increase in mortality with the new variant. Perhaps you could discuss this. Might you allude to the absolute risks here to help readers put results into perspective? The absolute risks are small (0.18% in the S+ group), and that is w/o the <30 yrs fraction. Perhaps these numbers should be in the abstract. Although you've included 109,545 patients, only 272 died.

* Might you also make the point that increased transmissibility might be more worrisome than increases in mortality with this variant? It's the spread that is the main problem with this VOC. A small increase in fatality might be considered "noise" at a population level.

Response:
These are important points.
We have expanded the discussion and conclusions to cover them, and included a sentence in the abstract. Also, we have expanded the first paragraph of the discussion to place the results in context. At a population level, any increase in death rate might be dwarfed by an increase in transmissibility, however at an individual-level measuring the IFR remains key for clinical decision making.

In the abstract:
"*In this comparatively low risk group, this represents an increase from 2.5 to 4.1 deaths per 1000 detected cases.*"

In the discussion:
"VOC-202012/1 infections (as measured by S-gene negativity) are associated with an elevated risk of death ($p<0.001$) in people testing positive for COVID-19 in the community. The increased hazard ratio (HR) between 1.32 and 2.04 over and above other variants, translates to a 32% to 104% increase in mortality, with the most probable HR estimate of 1.64, or a 64% increase. However, the absolute risk of death in this group of community-identified cases remains low, rising from 2.5 to 4.1 deaths per 1000 cases."

And in the conclusion:
"The resulting number of deaths will scale linearly with the proportion of cases infected with the new variant. Other analyses have indicated that the new variant is also associated with increased transmissibility, which would lead to a potentially exponential increase in the resulting number of deaths [12]."

* Can you clarify how you obtained mortality data?

Response:

We obtained the data from Public Health England under a data sharing agreement between PHE and our institutions. We have included a “data availability” statement at the start of the Methods section.

“Data availability
Data were collected by Public Health England (PHE) into a centralised database detailing the type of test performed and the results. PHE linked the line list data to hospital outcome and removed identifying information. PHE provided anonymised data to SPI-M contributors as part of the COVID-19 response under a data sharing agreement between PHE and the authors’ institutions.”

We also included the following expanded description of the death data:

“The deaths line list enumerates deaths in both hospital and community settings, and captures all deaths within 28 days of a positive COVID-19 test, following the Public Health England definition of death as ‘a death in a person with a laboratory-confirmed positive COVID-19 test and died within (equal to or less than) 28 days of the first positive specimen date’ [16]. Maintained by Public Health England, it represents the most timely and complete record of deaths due to COVID-19 in England. “

* The literature suggests a case fatality rate of somewhere between 0.8 and 1.2%, depending on hospital capacity. But here in a sample of adults 30 and up, the case fatality rate (within 28 days) appears to be 0.24%, 0.31% in the S-gene negative group and 0.18% in the S-gene positive group. Since you are examining 28 days within testing positive, and it takes anywhere from 5-10 days to present with symptoms, perhaps the followup period is too short? We also don't know whether their disease was more severe with respect to the need for acute care/hospitalization, etc.

Response:

The difference between our estimates and the 0.8%-1.2% estimate is due to the fact that our analysis relates to pillar 2 data, i.e. persons tested in the community who are at much lower risk of death than cases diagnosed in hospital. It is likely that 28 days is not long enough (from the PHE review 88% of deaths occurred within 28 days and we show in the supplement that the risk of death occurs mainly after day 14), although this is not the cause of the lower estimates.

* Please reconsider some of the terminology used in the paper. You say this is a matched cohort study but you used incidence density sampling - this is probably not the correct term as it is used for nested case-control studies. The terms "Cases" and "Controls" are also misused here; these two groups should be labelled as Exposed and Non-exposed groups, respectively.

Response:

We have updated this terminology as suggested. Where the categories “exposed” and “non-exposed” were not sufficiently clear we have used “S-gene negative cases” versus “S-gene positive cases”

* There is hardly any detail provided on the outcome ascertainment, censoring, and loss-to-follow-up data. We need to know, since a matched cohort study will prevent confounding (from the matched variables) given there was no (differential) loss-to-follow-up.

Response:

We obtain the data from PHE, who provide a technical report on the data processing (https://www.gov.uk/government/publications/phe-data-series-on-deaths-in-people-with-covid-19-technical-summary).

We have investigated censoring and loss-to-follow up further, although there is no obvious reason why cases infected with one variant should be more likely to be lost than the others - we show the distribution of times to censoring for each variant in supplementary figure S3. We now follow cases until 12th Feb 2020, thereby reducing censoring for the cases diagnosed at the end of December.

We have added the following section to the methods:

“Cases were followed up for 28 days following infection or until 12 February 2020, after which point cases that had no record of death were censored. In these data over 50% of COVID-19 related deaths are reported within 3 days and over 95% within 14 days (more details available in the supplementary materials). There are no differences in the reporting patterns for S-gene negative, and S-gene positive cases. The deaths line list is constructed from multiple sources as the gold standard list of COVID-19 related mortality in England, and ultimately will include all deaths where COVID-19 is mentioned on the death certificate. As a result we anticipate there are no COVID-19 deaths lost to follow up.”

And updated the results section:

“Every case was followed up for a minimum of 14 days, and over 85% of the cases were followed for the whole 28 day period.”

* The test eligibility was up to 29th of January. Does it mean that some people were only followed-up for just a few days (or even no follow-up at all)? This can be potentially quite problematic.

Response:

As described in the previous response, we have now added data up until 12 February 2020, thereby including a minimum of 14 days follow-up and 85% of cases are followed for the full 28 days. We describe this in the paper and graphically in the supplementary materials (figure S3).

* It was also not clear how 50 replicates were handled in the statistical analysis. Does it not artificially inflate the sample size?
* We are not clear how you integrated the multiple replicates and that might need some added explanation (possibly as an Appendix). Similarly, some further explanation of the creation of the replicates might be useful.

Response:

To clarify we used the 50 replicates to create 50 bootstrap estimates to test the impact of the matching procedure. Some cases had multiple matches in the control group, therefore the multiple replicates ensured that the estimates weren't contingent on a given match. We have clarified this in the methods section ('Data processing' paragraph 4) and the supplementary materials, in which we explain the matching process, replicates creation and combination of results.

"Some cases matched multiple controls and vice versa, so we sampled the cases and controls randomly within our framework to generate 50 replicates, ensuring no case or control was present more than once in each replicate. All analyses were conducted on each replicate as a separate sample and the results combined by combining the beta-coefficient estimates as a mixture of normal distributions, and calculated combination mean and confidence intervals numerically from the mixture distribution."

* There was a cross-over of the survival curves. Was the proportional hazards assumption violated?

Response:
The reviewer is correct that there is a violation of the proportional hazards assumption as seen in the Kaplan Meier figure which has a different pattern for the first 14 days compared to the second 14 days.

Investigating this, we can compensate for this non proportionality by considering these two time periods separately. Rather than confuse the main message of the paper we have conducted this analysis in the supplementary materials but acknowledge it more explicitly in the Results section text:

"The rate of death of S-gene negative and S-gene positive cases over time diverges after 14 days, shown in figure 2. This is graphical evidence that the hazard ratio is not constant over time, and as such this violates the proportional hazards assumption of the Cox model. This was investigated further (details presented in the supplementary materials) and this violation may be corrected by considering the hazard ratio in days 0-14 of follow up, versus that in days 15-28. The hazard ratio in the first period was not significantly elevated, but in days 15-28 the hazard ratio is found to be 2.40 (95% CI 1.66-3.47)."

* Please justify the 28-day time frame for the outcome.

Response:

The 28-day definition was developed by Public Health England, and is the longest follow up that can be justified with the current data. Therefore we used this cut off in order to be consistent with published data. In PHE's review, it was found that 88% of deaths occurred within 28 days, and that 46% of deaths excluded by the 28-day time limit did not have COVID on the death certificate.

* Finally, the mortality rate does appear considerably low compared to current levels (closer to 2% and not 0.2%). This would at least need a discussion. It's unclear how this would have affected your comparison. There does not seem to be a differential follow-up that would explain this as you are matching by spatio/temporal issues.
If the new strain is more common in those with multiple comorbidities, this could explain it, however, this might in itself be important and would require a note in the limitation section.

Response:
The low mortality of this group is reflecting the nature of the community diagnosis of COVID-19, resulting from the fact that S gene data is only available from Pillar 2. We have highlighted this fact throughout the manuscript, in the abstract, discussion and in the methods. Supplementary figure S1 shows the age distribution for Pillar 2 cases compared to the generation population, which demonstrates the tendency towards younger age groups for patients diagnosed in Pillar 2.

Paragraph 4, Methods:
"Antigen swab tests in the UK are carried out through two routes: Pillar 1 represents National Health Service (NHS) testing for those with a clinical need and healthcare workers; and Pillar 2 represents community testing of symptomatic individuals. Community based COVID-19 diagnoses are generally in a younger population with less severe disease than hospital based COVID-19 diagnoses, as elderly or severe cases tend to present directly to hospital (for more discussion see the supplementary materials)."

Paragraph 3, Discussion:
"This is a community- based study. We do not have information about the S-gene status of patients in hospitals. The community-based testing (Pillar 2) in this dataset covers a younger age group and hence represents less severe disease than cases detected through hospital- based testing (Pillar 1). In cases detected in the community, death remains a comparatively rare outcome, compared to in-hospital identified cases"

* Please better explain, for non-UK readers, what is meant by Pillar 1 and 2 testing and the populations that will be captured by each.

Response:
Thanks, we have included an explanation of Pillar 1 and Pillar 2 in the Methods section under "Inclusion Criteria", as well as replacing the term "Pillar 2 tests" by "community tests" throughout the manuscript.

"Antigen swab tests in the UK are carried out through two routes: Pillar 1 represents National Health Service (NHS) testing for those with a clinical need and healthcare workers; and Pillar 2 represents community testing of symptomatic individuals."

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how and where you have dealt with them in the paper.

Comments from Reviewers

Reviewer: 1

Comments:
"Increased hazard of mortality in cases compatible with SARS-CoV-2 variant of concern 202012/1 – a matched cohort study" compares rates of death in people with PCR tests exhibiting S-gene target failure (SGTF) with those without SGTF. The authors found that SGTF cases are significantly more likely to die within 28 days of diagnosis, after controlling for a number of factor. This is an important finding with timely and meaningful public health implications. The analysis is thorough and well-supported.

We thank the reviewer for careful reading of our manuscript.

Questions and recommendations:

The methods section says that participants who were diagnosed as late as January 29th were included in the analysis. Final disposition would not yet be available for most cases diagnosed in January. Can the authors clarify how recently-diagnosed cases were handled in the analysis?

Response:
We updated the analysis to include later available data, following cases until 12 February 2021, so that we have a minimum of 14 days follow-up per person. This led to an increase in the number of matched pairs and fewer cases rejected due to data quality problems. The additional follow up also now includes more deaths in the sample, and fewer censored cases. Figure S3 in the supplementary material shows the length of follow-up per case, which shows that 85% of cases are followed for the full 28 days.

"Cases were followed up for 28 days following infection or until 12 February 2020, after which point cases that had no record of death were censored. In these data over 50% of COVID-19 related deaths are reported within 3 days of date of death and over 95% within 14 days [15] (more details available in the supplementary materials). Every case was followed up for a minimum of 14 days after the positive test result, and over 85% of the cases were followed for the whole 28 day period. There are no differences in the reporting patterns for S-gene negative, and S-gene positive cases. The death line list is constructed from multiple sources as the gold standard list of COVID-19 related mortality in England, and ultimately will include all deaths where COVID-19 is mentioned on the death certificate."

Although Public Health England has reported that SGTF is highly correlated with the B.1.1.7 lineage in the UK, that has not always been the case in the United States [1]. I recommend clarifying the robustness of SGTF as a marker for the VOC in your dataset, to benefit international readers.

Response:

We have included the following text in the methods, and cited the Helix dashboard to clarify the US situation.

"SGTF cases have subsequently been used as a proxy to track the progression of this variant in the UK [7–9,2], which is possible because of the strong association in the UK between B.1.1.7 and S-gene negative cases".

"Pillar 2" is UK-specific jargon. I recommend clarifying what constitutes a Pillar 2 test (beyond what is stated on page 4 line 12, which I did not find clarifying) and use a more common term throughout.

Response:

Thanks. We have replaced Pillar 2 tests with community-based tests throughout the manuscript and added a further description in the methods in the "Inclusion Criteria" section.

The text asserts that the VOC was first detected in the UK in December of 2020. Public Health England situation reports suggest that detection was as early as September [3].

Response:
It is correct that B.1.1.7 was identified in October, but it was only identified as "a variant of concern" in December. We haved update the text to:

"A new lineage of the SARS-CoV-2 virus (named B.1.1.7) was identified from genomic sequencing of cases in the South East of England in early October 2020. It was identified as a variant of concern of the SARS-CoV-2 virus [1] (VOC-202012/1 - 'new variant') in December 2020 by Public Health England."

There are a number of grammatically errors that should be fixed before publication.

Thanks, we have carefully checked the manuscript for errors.

[1] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957504/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5_England.pdf
[2] https://www.helix.com/pages/helix-covid-19-surveillance-dashboard
[3] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/947048/Technical_Briefing_VOC_SH_NJL2_SH2.pdf

Please enter your name: Caitlin Rivers
Job Title: Assistant Professor
Institution: Johns Hopkins

Comments:
In this paper, the authors conduct a pair-matched analysis of community testing data and death data to estimate the relative mortality of the SARS-CoV-2 variant of concern in the UK. This is one of several analyses of these data that have been done by research groups, at the request of the government. The pair-matched approach enables the authors to match finely on several covariates, including age, locality, and time. The authors estimate a hazard ratio of 1.7, indicating an increase in mortality associated with this variant.
The question is clearly highly impactful, as it has important public health implications for the UK and the rest of the world. The pair-matched approach is sensible, and allows very fine matching, although at the expense of unmatched cases who then do not contribute to the analysis. The authors comment on this trade-off in the discussion. I provide comments on a few reservations I have. First, the methods are not fully described, and more precision will be needed throughout. Second, the logic of the investigation into sources of bias is not always clear. On the other hand, there are other types of sensitivity analyses or subanalyses not included that would help readers evaluate the robustness of the result. I provide several comments below.

We thank the reviewer for careful reading and helpful comments which we address below.

Major comments:
1) The authors do not discuss lags in death reporting and how these may affect recent cases. But it seems likely to impact the results as the analysis includes individuals whose first test was on Jan 29th, 2021, for an analysis written on Jan 31st, 2021. How was censoring defined? This should be explained very precisely as a key study method.

Response:

We have updated the analysis to include data up to 12 February 2021, while keeping the inclusion criterion that cases were diagnosed before 29 January 2021, thereby having a minimum of 14 days follow up per case (shown in detail in figure S3).

A detailed examination of reporting delays has been carried out PHE and other groups, and we have included a reference to a pre-print that discusses reporting delays to deaths in further detail (Seaman S, Samartsidis P, Kall M, et al. Nowcasting CoVID-19 Deaths in England by Age and Region. medRxiv 2020;:2020.09.15.20194209. doi:10.1101/2020.09.15.20194209).

"Cases were followed up for 28 days following infection or until the 12th Feb 2020, after which point cases that had no record of death were censored. In these data over 50% of COVID-19 related deaths are reported within 3 days of date of death and over 95% within 14 days [15] (more details available in the supplementary materials). Every case was followed up for a minimum of 14 days after the positive test result, and over 85% of the cases were followed for the whole 28 day period. There are no

differences in the reporting patterns for S-gene negative, and S-gene positive cases. The deaths line list is constructed from multiple sources as the gold standard list of COVID-19 related mortality in England, and ultimately will include all deaths where COVID-19 is mentioned on the death certificate."

2) The authors describe how they identified 50 sets of matched pairs, to generate a more robust result that integrates more of the available data, but they do not describe how this was accommodated in the modeling, point estimation, and confidence interval estimation. This should be explained precisely.

Response:

We have expanded the methods section for clarification, and include more detail in the supplementary information.

"All analyses were conducted on each replicate as a separate sample and the results combined by combining the beta-coefficient estimates as a mixture of normal distributions, and calculating combination mean and confidence intervals numerically from the mixture distribution (further discussion of this topic is found in the supplementary materials)."

3) Can the authors explain why CT value would be regarded as a potential source of bias? Given that CT could lie along the causal pathway, it may be more interpretable if the analysis is presented as an effect that is not exclusively explained by CT. But it is worth pointing out that adjusting for N gene CT undermines the matching procedure employed.

Response:

We don't regard CT value as a potential source of bias, but suggest it could be interpreted as such - we have clarified the text in the results section of the paper. We have added the following sentence:

"The higher viral load of S-gene negative cases could be the biological result of the VOC202012/1 mutations and be the cause of higher mortality, but alternatively it could be an indication of the timing of testing, with S-gene negative cases presenting at peak infectiousness, for some as yet unknown reason."

Given the uncertainty around the role of viral load in the causal pathway, matching patients by CT values may have introduced its own bias; this we chose to allow it as a free variable but investigate it as a covariate.

From Shinozaki paper referenced, "when additional confounders are adjusted in the analyses, such cancellation breaks down and ignoring matching variables results in biased estimates."

Response:

The Shinozaki method was used only for cross checking the unadjusted models, and we did not include the results in the main text in the interest of brevity. It is not the method used for the reported estimates which are standard Cox models. We removed reference to this paper for clarity and although we did cross reference our unadjusted estimates with this method we do not present them in the paper.

Similarly, why do the authors not adjust for age in the N gene CT model, where age had been shown in the prior model to help explain hazard.

Response:
This analysis was conducted but not originally included. We have updated the CT model to include age which was found to have an independent effect on N Gene CT. This does not change the main result.

4) Can the authors explain why they control for age and not the other matched covariates in their model? It seems that these could improve the precision of the model by removing sources of variability in hazard. Did the authors explore more flexible models for age than a single linear term?

Response:

We controlled for age as it was the variable that we allowed the most tolerance for during the case matching. We have now included a broader range of models including more covariates into the supplementary materials and the results remain robust to covariate inclusion.
We also investigated including the age as a non linear spline but the resulting spline was essentially a straight line, so we reverted to a simple linear term. We have included this analysis in the supplementary material.

5) Table 2. “Hazard rate” should be “Hazard ratio.” Figure 4A-C. “Hazard rate” should be “Hazard ratio.” Similarly, Figure 4 caption. “Slightly lower estimates of hazard” should be “slightly lower estimates of the hazard ratio.” Page 7. “Shows a reduction in the overall hazard of S-gene negativity to 1.4” should be “shows a reduction in the overall hazard ratio of S-gene negativity to 1.4.” Page 7. “Figure 4 shows the estimates of hazards related to alternating those assumptions.” And so on.

Response:

We have checked the text and now use the term Hazard Ratio consistently.

6) Section titled Sensitivity analysis. Inadequate detail is provided to the reader, who may not be familiar with the author’s preliminary analysis. I suggest providing the prior estimate for context, an accounting of how much more data were made available, and any changes in definitions.

Response:

We have restructured the discussion section to include our previous work with other estimates

"In preliminary work on this topic we estimated the hazard ratio to be 1.91 (95% CI 1.35 - 2.71), which is marginally higher than the estimate presented here with compatible uncertainty [25,26]. This was based on 94 S-gene negative deaths and 49 S-gene positive deaths in 66,208 less strictly matched pairs. As the situation has unfolded and more data has become available we have been able to narrow our tolerance for mismatches, and increase the proportion of cases with complete follow up, to get more accurate central estimates"

7) Decreasing the CT threshold increases the number of equivocal cases, but I would encourage the authors to point out that this would preferentially target the S-gene positive cases (making them more equivocal) as these have systematically higher CT. So in that way, it has the potential to induce bias.

Response:

Thanks - that is a good observation. We have added the following text to the methods section:

"Given the fact that CT values are generally higher is S-gene positive cases, reducing the CT threshold will tend to reclassify more mild S-positive cases than mild S-negative cases as equivocal and subsequently exclude them from the analysis. This could explain the small reduction in hazard ratio associated with reducing the CT threshold."

8) In terms of assessing biological plausibility of increased severity, it would be helpful to see how this hazard ratio holds up across subgroups. Analyses that examine the robustness of the finding across subgroups would provide greater confidence than some of the sensitivity analyses currently included.

Response:

We agree that this would be an interesting consideration, but are hampered by the sparsity of the data. Once further data become available to allow the study of subgroup hazard, it will be of substantial value. We have included a model that includes more covariates in the supplementary materials, which gives the same answer, but it is not powered to detect variation in the subgroups. The discussion and conclusions reflect this point.

Minor comments:
9) Figure 1 and Text. I assume it is rounding of an average, but people will notice that 54,773*2 != 109,545.

Response:

Thank you. We have edited figure 1 and text for clarity although with the inclusion of more follow up data the figures have changed.

10) Kaplan Meier curve needs a more informative x-axis, e.g. time from first positive test.

Response:

Thank you. We have edited Figure 2

11) Table 2. Suggest reporting hazard ratios for age and CT value as 10 unit changes, or including more significant digits. 0.9 (0.9 – 0.9) has a very strange appearance.

Response:

Thank you - we have made changes to the presentation of the models in Table 2.

12) P6 Line 49. Admission is presumably hospital admission, but please clarify in the text. Hospital admission data are not described in the study methods. Are these linked to the death data?

Response:

We have replaced “admission” with “hospitalisation” for clarity. The hospital admission information is only available for patients that die and is included in the data by PHE. We have added the following to the methods:

“... The deaths line list also contains some details about the timing of hospitalisation in those people that died.”

13) Figure 3A and the associated text. The authors should make it clearer to readers that this analysis is restricted to individuals who ultimately died. How does this analysis handle individuals who died but were not admitted to the hospital?
14) Figure 3A. The authors should clarify what bias this analysis is meant to signal.

Response:

We have updated text to include:

“One possibility for bias could be a difference in the timing of the presentation of S-gene negative and S-gene positive infections for testing, with for example S-gene positive cases presenting earlier, and thus appearing to progress slower. We only have hospital admission data for patients who ultimately died but in this there is no evidence for asymmetric delays in time from test to hospital admission as shown in figure 3 panel A. This has also been investigated by the Office of National Statistics who find that S-gene negative patients are more likely to present earlier for testing [23]. “

15) It is unclear why Figure 3B would signal a source of bias. Instead, as the authors note, it reflects when circulation of the new variant took off, enabling enough pairs to be matched.

This seems like this information would be better suited for Table 1, by adding a row for each month, providing a summary of the data.

Response:

We believe a figure is able to describe the extent of the rise in the number of both S-positive and S-negative cases in a more compact way. Since large numbers of cases in hospitals, independent of which strain they are infected with, could be associated with increased mortality, we believe the time period for the analysis could be a source of potential bias.

We have also added the following text:

"As the degree of hospital load in the UK varied considerably over the study period it is important to note that there is no avoidable temporal clustering in the matched cases."

16) Figure 4D. Pairwise bias not clearly defined.

Response:

We altered the text referring to panel D and figure caption:

"Because of the change in prevalence from predominantly S-gene positive to predominantly S-gene negative, increasing the tolerance to sample date mismatch leads to a systematic "pairwise bias" in the dates of the original positive test result (Panel D) with S-gene negative cases generally being identified after S-gene positive controls."

Figure caption:

"... because the prevalence of cases and controls changes over time, it also results in increasing systematic bias between pairs of cases and controls (labelled "pairwise bias" in panel D) ..."

17) Discussion. I would suggest more cautious language than "VOC-202012/1 infections… lead to an elevated risk of death." Was associated with.

Response:

Thank you - we have checked the text and used more cautious language.

18) The authors completed the STROBE checklist for a case-control study, despite this being a cohort analysis.

Response:

This has been updated