

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) | Diagnostic Accuracy of X-ray versus CT in COVID-19: A Propensity Matched Database Study |
| AUTHORS | Borakati, Aditya; Perera, Adrian; Johnson, James; Sood, Tara |

VERSION 1 – REVIEW

| | |
|------------------------|--|
| REVIEWER | Zhonghua Sun Curtin University, Australia |
| REVIEW RETURNED | 31-Jul-2020 |

| | |
|-------------------------|---|
| GENERAL COMMENTS | <p>This is a nicely written manuscript further reporting the usefulness of chest CT in the diagnosis of COVID-19 with high sensitivity when compared to chest x-ray. Although it is a single centre study, it was performed on a large number of cases (1198 patients) presented to the ED. As authors indicated, data were analysed using comprehensive statistical analyses. I do not have main concerns of accepting it for publication. Some minor comments are listed below:</p> <ol style="list-style-type: none">1. Methods: please provide CT scanning protocols, including CTPA as this affects the resultant radiation dose.2. There are some short paragraphs only containing one sentence, mainly in the Discussion and I would suggest authors try to combine them into one paragraph. We should avoid having only one sentence in a paragraph. |
|-------------------------|---|

| | |
|------------------------|--|
| REVIEWER | Ming Yen Ng The University of Hong Kong, Hong Kong SAR, China Received funds from Bayer AG and Circle Cardiovascular Imaging |
| REVIEW RETURNED | 08-Aug-2020 |

| | |
|-------------------------|--|
| GENERAL COMMENTS | <p>Borakati et al, have undertaken a retrospective study with 1198 patients who attended the Emergency department and had CXR and RT-PCR testing performed. The objective was to determine the diagnostic accuracy of CXR and also CT where CT was available in these patients.</p> <p>This is a large study and authors should be applauded for the substantial work undertaken. However, there are some significant issues which need to be dealt with in order for this manuscript to be publishable.</p> <p>Major Comments:</p> <ol style="list-style-type: none">1. The abstract and manuscript conclusion in its current state goes further than the research question with comments that the study was not designed to answer. This is a study on diagnostic accuracy in the situation of the emergency department. It is not possible to draw a conclusion on the suitability of utilising CT preferentially over CXR. |
|-------------------------|--|

| | |
|--|--|
| | <p>This is a highly complex situation which depends on each centres CT access, infection control policies, access to rapid RT-PCR kits, etc. One reason given by the authors is In the discussion section in which the advantage of CT is the "rapid turnaround". This at a minimum requires a reference but in order to make this comment, assessing the time taken for a CT to be performed in relation to CXR should be done. Overall, this conclusion needs to be adjusted to focus on the answering the research question which is the diagnostic accuracy of CXR and CT.</p> <p>2. Related to the above point, the authors should comment in the discussion about the role of RT-PCR. The drive is to use rapid RT-PCR testing kits. This would have significant implications on the impetus to use CT over CXR.</p> <p>3. One major confounder not mentioned in this study is selection bias. Out of 1198 patients with CXR, there were only 302 patients with CT (ie. 25% of patients). I wonder if the patients who had CT were more unwell or likely suspected to have pulmonary embolus thus the likelihood of CT finding abnormalities would be higher. This needs to be more adequately discussed. Furthermore, 72% (87/319) of patients having CT were COVID-19 positive whilst 52% of CXR finally included were COVID-19 positive. It is also noticeable that 449/1022 COVID-19 negative CXR patients were excluded for lack of RT-PCR tests. This is another source of bias.</p> <p>4. The criteria for diagnosing COVID-19 needs to be clarified. It is not entirely clear whether one needs just a single RT-PCR result was required to be classed as positive for COVID-19. What about patients with strong clinical suspicion but negative RT-PCR? In order to make sense of the diagnostic accuracy data. A comment on the accuracy of RT-PCR in their centre would also be useful.</p> <p>5. Discussion section - correlation between CXR and laboratory tests, I don't think adds to the discussion and should be removed. Although the authors did not find a significant correlation, this would for me suggest that actually the combination of blood tests and CXR would improve diagnostic accuracy as the two types of investigations are looking for the disease from different vantage points.</p> <p>6. How much of the data had missing values? A sense of this would help the readers understand how much of the data was imputed. Statistical review is required for this.</p> <p>7. Discussion - Mention about infection control policies and access to CT scanners is an important discussion point and was one reason for recommendations against the use of CT. The disparity between some nations using CT versus CXR has been access to CT scanners and decisions on infection control China opted for CT due to the large amount of scanners available in their respective hospitals allowing them to dedicate their scanners for COVID-19. This was not so feasible in other nations.</p> <p>8. What proportion of the CXRs were PA films? This will likely have an impact on diagnostic accuracy.</p> <p>9. If indeterminate findings on CXR are classed as positive for COVID-19, the sensitivity of CT and CXR diminishes considerably (0.85 vs 0.80) with specificity of (0.50 vs 0.40). Furthermore, a</p> |
|--|--|

| | |
|-------------------------|---|
| | <p>normal CXR does not exclude the diagnosis of COVID-19 as this is a well known finding in COVID-19 patients. So a sensitivity analysis of CXR including normal CXR would like increase CXR's sensitivity further. For the manuscript, providing these results would actually produce a more balanced presentation of the data allowing readers to decide for themselves what degree of sensitivity/ specificity is acceptable for their units.</p> <p>10. One limitation not mentioned, is the prevalence of COVID-19 coming to the hospital at the time of this study. This will have a significant impact on the generalisability of the diagnostic accuracy of any test.</p> |
| REVIEWER | Lina Zhang The First Hospital of China Medical University. China |
| REVIEW RETURNED | 20-Aug-2020 |
| GENERAL COMMENTS | <p>1. Covid-19 is a rapidly changing disease. If the CT and X-ray diagnostic ability are to be compared, the two examinations should be conducted at approximately the same time. From the description of the article, I am not sure whether the two examinations are in the same period. If so, it is suggested to add a flow chart to make examination date clear; if not, please carefully consider the comparison between the two examinations .</p> <p>2.What kind of patients underwent CT examination during the 30 days? Were those patients who underwent CT examination more seriously? Is it possible that the difference between patients with and without CT might affect the results?</p> <p>3.In Figure 1, why does "intermediate" appear in the "CT / X-ray negative" column? It is suggested to add the necessary explanations. such as What does the "unmatched" mean?</p> <p>4.There was discussion about PEs, but no data was found in the results. Please keep the results consistent with the discussion.</p> <p>5.Inter-rater reliability of imaging reports is suggested.</p> |

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name
Zhonghua Sun

Institution and Country
Curtin University, Australia

Please state any competing interests or state 'None declared':
None

Please leave your comments for the authors below

This is a nicely written manuscript further reporting the usefulness of chest CT in the diagnosis of COVID-19 with high sensitivity when compared to chest x-ray. Although it is a single centre study, it was performed on a large number of cases (1198 patients) presented to the ED. As authors indicated, data were analysed using comprehensive statistical analyses. I do not have main concerns of accepting it for publication. Some minor comments are listed below:

1. Methods: please provide CT scanning protocols, including CTPA as this affects the resultant

radiation dose.

2. There are some short paragraphs only containing one sentence, mainly in the Discussion and I would suggest authors try to combine them into one paragraph. We should avoid having only one sentence in a paragraph.

Many thanks for your positive comments, we have added the CT scanning protocols and amended the sentences.

Reviewer: 2
Reviewer Name
Ming Yen Ng

Institution and Country
The University of Hong Kong, Hong Kong SAR, China

Please state any competing interests or state 'None declared':
Received funds from Bayer AG and Circle Cardiovascular Imaging

Please leave your comments for the authors below

Borakati et al, have undertaken a retrospective study with 1198 patients who attended the Emergency department and had CXR and RT-PCR testing performed. The objective was to determine the diagnostic accuracy of CXR and also CT where CT was available in these patients.

This is a large study and authors should be applauded for the substantial work undertaken. However, there are some significant issues which need to be dealt with in order for this manuscript to be publishable.

Major Comments:

1. The abstract and manuscript conclusion in its current state goes further than the research question with comments that the study was not designed to answer. This is a study on diagnostic accuracy in the situation of the emergency department. It is not possible to draw a conclusion on the suitability of utilising CT preferentially over CXR. This is a highly complex situation which depends on each centres CT access, infection control policies, access to rapid RT-PCR kits, etc. One reason given by the authors is In the discussion section in which the advantage of CT is the "rapid turnaround". This at a minimum requires a reference but in order to make this comment, assessing the time taken for a CT to be performed in relation to CXR should be done. Overall, this conclusion needs to be adjusted to focus on the answering the research question which is the diagnostic accuracy of CXR and CT.

We have stated that CT should only be used where capacity allows in every sentence where we advocate it's use. We also mention that rapid RT-PCR kits are more widely available in the discussion and how this may obviate CT, we have added a section on infection control.

In regards to rapid turnaround of CT we mean in comparison to RT-PCR, not in comparison to X-ray, which will of course be even faster.

We have added a reference to the time taken for CT, which takes seconds for a non-contrast CT or 1-2 minutes for a contrast CT. PCR will take a minimum of 1 hour to perform ignoring logistics.

2. Related to the above point, the authors should comment in the discussion about the role of RT-PCR. The drive is to use rapid RT-PCR testing kits. This would have significant implications on the impetus to use CT over CXR.

We have expanded on this, rapid RT-PCR is very limited in capacity in many countries and in any case will take at least 1 hour to perform, ignoring logistics, which is significantly longer than the time taken to conduct a non contrast or contrast chest CT (on the order of seconds to minutes).

3. One major confounder not mentioned in this study is selection bias. Out of 1198 patients with CXR, there were only 302 patients with CT (ie. 25% of patients). I wonder if the patients who had CT were more unwell or likely suspected to have pulmonary embolus thus the likelihood of CT finding

abnormalities would be higher. This needs to be more adequately discussed. Furthermore, 72% (87/319) of patients having CT were COVID-19 positive whilst 52% of CXR finally included were COVID-19 positive. It is also noticeable that 449/1022 COVID-19 negative CXR patients were excluded for lack of RT-PCR tests. This is another source of bias.

The proportion of images reported as positive for CXR and CT, are 52% and 72% respectively, not those who tested positive by RT-PCR swab. Given that CT has greater resolution and sensitivity than X-ray this is not surprising that a higher percentage were reported positive. There were (548/860) 63.7% RT-PCR positive patients in propensity matched X-ray patients. In the CT group with RT-PCR swabs there were 63.2% (191/302) who tested positive by RT-, therefore, the proportion testing positive in both CT and X-ray groups is similar.

It is correct that a higher proportion of patients with X-ray were excluded for not having RT-PCR versus CT however. Unfortunately testing capacity was very limited in the UK at the time of our study. We have expanded on this further and added these to the limitations section. However, this bias only increases the observed sensitivity and specificity of CXR in our study relative to CT. Without this bias these values would be lower, therefore in reality CT is likely to be even more sensitive than CXR than as reported in our study.

Ideally an RCT or other prospective study is needed to ensure all patients receive all tests and remove biases, however, this would be difficult to conduct in practice.

With regards to the patients undergoing CT we have further expanded on the selection bias.

4. The criteria for diagnosing COVID-19 needs to be clarified. It is not entirely clear whether one needs just a single RT-PCR result was required to be classed as positive for COVID-19. What about patients with strong clinical suspicion but negative RT-PCR? In order to make sense of the diagnostic accuracy data. A comment on the accuracy of RT-PCR in their centre would also be useful.

All RT-PCR swabs taken in 30 days of initial attendance to ED were used as stated in the methods. Where any swab was positive on a patient this was taken to be positive overall.

There were 810 RT-PCR swabs conducted on 435 patients who tested negative overall across the 30 days. The repeated testing on the negative patients suggests there was clinical suspicion of COVID-19, however, we have no data on whether they were clinically suspected of having COVID-19 from documentation.

We cannot comment on the accuracy of RT-PCR here as there is no reference standard to compare it's accuracy to. Other studies have used CT as the reference standard, in which case the sensitivity of RT-PCR here would be $162/217 = 74.7\%$.

RT-PCR and swabbing at our centre was performed according to national and World Health Organisation protocols, we have no reason to suspect accuracy of RT-PCR was any different at our centre versus anywhere else.

5. Discussion section - correlation between CXR and laboratory tests, I don't think adds to the discussion and should be removed. Although the authors did not find a significant correlation, this would for me suggest that actually the combination of blood tests and CXR would improve diagnostic accuracy as the two types of investigations are looking for the disease from different vantage points.

On multivariable analysis only neutrophils and CRP were independently associated with COVID-19 diagnosis. These are both non-specific for COVID-19 and while, these in addition to the clinical picture and imaging contribute to the clinical diagnosis of COVID-19, we have not combined these as a measure and analysed these. Given that CT is more sensitive there is potentially the argument that CT and blood results would be superior to CXR and blood results. We have added to the discussion about combining these modalities for diagnosis however.

We believe that correlating imaging findings with laboratory and clinical parameters is important, as it allows clinicians to identify whether imaging is a valuable marker of severity and prognosis in COVID-19.

6. How much of the data had missing values? A sense of this would help the readers understand how much of the data was imputed. Statistical review is required for this.

The proportion of missing values for each variable is shown in table 1. Variables with more than 50% missing were removed for analyses and not imputed. The maximum proportion of missing was 24% in one variable after this removal.

This has been reviewed by the Department of Statistical Science at UCL and analyses have been confirmed as robust including imputation as stated in the title page. Our analysis code is published online as well as enclosed as a supplementary file and the link is available in the manuscript: [10.6084/m9.figshare.12674099](https://doi.org/10.6084/m9.figshare.12674099)

7. Discussion - Mention about infection control policies and access to CT scanners is an important discussion point and was one reason for recommendations against the use of CT. The disparity between some nations using CT versus CXR has been access to CT scanners and decisions on infection control China opted for CT due to the large amount of scanners available in their respective hospitals allowing them to dedicate their scanners for COVID-19. This was not so feasible in other nations.

We have mentioned the caveat of capacity in each sentence where we have advocated CT. We have added to the discussion infection control policy.

8. What proportion of the CXRs were PA films? This will likely have an impact on diagnostic accuracy.

All CXRs were PA.

9. If indeterminate findings on CXR are classed as positive for COVID-19, the sensitivity of CT and CXR diminishes considerably (0.85 vs 0.80) with specificity of (0.50 vs 0.40). Furthermore, a normal CXR does not exclude the diagnosis of COVID-19 as this is a well known finding in COVID-19 patients. So a sensitivity analysis of CXR including normal CXR would like increase CXR's sensitivity further. For the manuscript, providing these results would actually produce a more balanced presentation of the data allowing readers to decide for themselves what degree of sensitivity/specificity is acceptable for their units.

Sensitivity of both modalities increased as described in the results, when indeterminates were taken as positive:

"Taking X-rays reported as indeterminate as positive increased the sensitivity of CXR to 0.80 (95% CI 0.77-0.84), however reduced specificity to 0.40 (95% CI 0.35-0.46). When CT scans reported as indeterminate are also considered positive the sensitivity of CT increased to 0.93 (95% CI 0.89-0.96), whilst mean specificity reduced to 0.37 (95% CI 0.28-0.47),"

We are not sure the value of adding images reported as negative for COVID or alternate pathology as a positive. This would essentially mean all X-rays are positive, this would of course mean that the sensitivity is 100%, however, the specificity would be much lower. Also, if essentially the imaging is always positive, there is no point in conducting the imaging as you know it is positive. We have done a sensitivity analysis with indeterminate images however.

We feel these comments support our conclusions. If it is well known that clear CXRs can still mean COVID-19 is present (and a very large proportion of these as the sensitivity is poor as in our study), it begs the question as to why CXRs are being performed for COVID-19 diagnostics as we know that they add little value. As above classing clear CXRs as positive, essentially renders CXR meaningless as they will always be positive in that case and cannot distinguish positive and negative cases.

10. One limitation not mentioned, is the prevalence of COVID-19 coming to the hospital at the time of

this study. This will have a significant impact on the generalisability of the diagnostic accuracy of any test.

We agree this will influence our results significantly. Unfortunately at the time of our study in the UK, mass testing was not available and consequently the prevalence of COVID-19 in the community was unknown (this appears to be the same in most countries). We have added some references to some estimates of the prevalence of COVID-19 in the UK however.

Reviewer: 3
Reviewer Name
Lina Zhang

Institution and Country
The First Hospital of China Medical University.
China

Please state any competing interests or state 'None declared':
NONE

Please leave your comments for the authors below

1. Covid-19 is a rapidly changing disease. If the CT and X-ray diagnostic ability are to be compared, the two examinations should be conducted at approximately the same time. From the description of the article, I am not sure whether the two examinations are in the same period. If so, it is suggested to add a flow chart to make examination date clear; if not, please carefully consider the comparison between the two examinations.

We agree and have stated this as a limitation. However, 30.1% of CTs were performed on the same day as CXR and the median time to CT was 4.5 days (see results), therefore the majority were performed at a similar time to CXR.

2. What kind of patients underwent CT examination during the 30 days? Were those patients who underwent CT examination more seriously? Is it possible that the difference between patients with and without CT might affect the results?

In our hospital, CT was only performed if there was suspicion of pulmonary embolus (D-dimer >5000 as per local guideline or clinical suspicion of PE, where CTPA was performed) or negative X-ray but ongoing clinical suspicion of COVID-19. It is a possibility that only more severe patients may have received CT and this is a limitation, we have expanded on this further in the discussion.

3. In Figure 1, why does "intermediate" appear in the "CT / X-ray negative" column? It is suggested to add the necessary explanations. such as What does the "unmatched" mean?

'Indeterminate' images were classed as negative. Only those reported as Classic COVID-19 as per the BSTI guidelines were considered positive in the main analyses. We considered indeterminate images as positive in a secondary analysis in the results also.

Unmatched refers to patients excluded as a result of propensity score matching. This is a statistical technique to minimise confounding in addition to multivariable regression that we have used.

4. There was discussion about PEs, but no data was found in the results. Please keep the results consistent with the discussion.

We have added the data to the results.

5. Inter-rater reliability of imaging reports is suggested.

Unfortunately this is not possible as each image was only reported by one consultant and we do not have permission to compare the accuracy of reporting for each consultant. This has been stated as a limitation however.

VERSION 2 – REVIEW

| | |
|------------------------|--|
| REVIEWER | Ming Yen Ng The University of Hong Kong Received funds from Bayer AG and Circle Cardiovascular Imaging |
| REVIEW RETURNED | 27-Sep-2020 |

| | |
|-------------------------|--|
| GENERAL COMMENTS | <p>The reviewers have responded to my comments and provided further information to help me better understand the manuscript. Nonetheless, I have further comments.</p> <ol style="list-style-type: none"> 1. The abstract's conclusion still goes beyond the research question/objective of comparing the diagnostic accuracy of CT and CXR. The study did not compare the "turnaround time" of CT vs CXR strategy for patients and as a result the meaning of turnaround time is unclear. The paper is a diagnostic study of CT versus CXR. Practically for most radiology departments cleaning time of the machine and ensuring non-COVID-19 patients are separated from suspected COVID-19 patients takes time. There is also the personal protective equipment that the radiology department staff will need to wear. These were not assessed in this study and therefore parts of the conclusion are inappropriate. Discussion of the potential of CT in the discussion is entirely reasonable but as a statement in the abstract, this is not appropriate and is misleading. Only the first sentence of stating that CT has better diagnostic accuracy is appropriate and supported by this study's results. Consider stating the conclusion as "Computed tomography has substantially improved diagnostic performance over CXR in COVID-19. CT should be considered in the initial assessment for suspected COVID-19 instead of CXR if capacity allows and balanced against radiation exposure risk." 2. Table 3, the terms "Apparent prevalence" and "True prevalence" should be explained as to what prevalence the authors are referring to. This will improve clarity of the table. 3. The diagnosis of patients with COVID-19 is still unclear in the manuscript. Furthermore, for RT-PCR results to be valid for 30 days from initial ED attendance leaves many questions about the separation of the cohort into COVID-19 positive and negative. This needs to be listed as a limitation. It is well documented that COVID-19 patients can have negative RT-PCR results at presentation. If it is possible that some patients in this study could be in this category, this should also be stated as a limitation. 4. As the sensitivity of the RT-PCR test is unknown, please add references to the sensitivity of RT-PCR from previously published studies. Consider the following BMJ article to reference: Jessica Watson, Penny F Whiting, John E Brush. Interpreting a covid-19 test result. BMJ 2020. https://www.bmj.com/content/bmj/369/bmj.m1808.full.pdf |
|-------------------------|--|

| | |
|--|---|
| | <p>5. As CXRs are all PA films, this should be stated in the methods section to allow the readers to understand that these are high quality CXRs unlike AP CXRs. Consider mentioning this in the discussion as this not the case in all centres worldwide where AP films are relatively common especially in unstable patients.</p> <p>6. To provide clarity of the CXR report gradings in table 2, it should state that this follows BSTI grading. Furthermore, the grades should be stated with the descriptions. For example, “Grade 0 – Alternative pathology”, “Grade 1 - No abnormalities”. In the results section a brief description of the number of patients with the gradings of the CXR and CT would provide further clarity to the paper.</p> <p>7. In Methods section under Diagnostic Accuracy Statistics (page 103 of 127 line 26), please include a reference to the BSTI COVID-19 CT reporting template.</p> <p>8. Under the strengths and limitations in the discussion section (page 111 of 127, line 34), the following line “This study is the largest conducted on imaging in the COVID-19 pandemic” needs to be amended. Please see the following references of larger studies conducted on COVID-19 imaging to help amend the sentence: i) G Herpe, M Lederlin, et al. Efficacy of Chest CT for COVID-19 Pneumonia in France. Radiology (2020) ii) K De Smet, et al. Diagnostic Performance of Chest CT for SARS-CoV-2 Infection in Individuals with or without COVID-19 Symptoms. Radiology (2020) iii) Ng M, Wan E, Wong H et al (2020) Development and Validation of Risk Prediction Models for COVID-19 Positivity in a Hospital Setting. International Journal of Infectious Diseases (2020)</p> <p>9. Under the the strengths and limitations in the discussion section (page 112 of 127, lines 28-32), the comment of “This is likely the case worldwide however as many asymptomatic patients remain undiagnosed and those with mild disease are encouraged to self isolate rather than see medical professionals and be tested” should be adjusted as different localities have adopted mass screening programmes of asymptomatic patients at different timepoints including the UK but on a smaller scale. Examples, include Wuhan, China around May 2020 and a link below to the University of Nottingham asymptomatic screening programme. https://www.nottingham.ac.uk/news/university-of-nottingham-to-develop-asymptomatic-testing-service-to-help-control-hidden-spread-of-covid-19</p> <p>10. Consider adjusting the statement in manuscript conclusion (page 114 of 127, line 10) to reflect the study findings and grammar “and should strongly be considered during the pandemic...” to “and should be considered ahead of chest x-ray in the initial assessment of COVID-19”</p> |
|--|---|

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Reviewer Name

Ming Yen Ng

Institution and Country

The University of Hong Kong

Please state any competing interests or state 'None declared':

Received funds from Bayer AG and Circle Cardiovascular Imaging

Please leave your comments for the authors below The reviewers have responded to my comments and provided further information to help me better understand the manuscript. Nonetheless, I have further comments.

1. The abstract's conclusion still goes beyond the research question/ objective of comparing the diagnostic accuracy of CT and CXR. The study did not compare the "turnaround time" of CT vs CXR strategy for patients and as a result the meaning of turnaround time is unclear. The paper is a diagnostic study of CT versus CXR. Practically for most radiology departments cleaning time of the machine and ensuring non-COVID-19 patients are separated from suspected COVID-19 patients takes time. There is also the personal protective equipment that the radiology department staff will need to wear. These were not assessed in this study and therefore parts of the conclusion are inappropriate. Discussion of the potential of CT in the discussion is entirely reasonable but as a statement in the abstract, this is not appropriate and is misleading. Only the first sentence of stating that CT has better diagnostic accuracy is appropriate and supported by this study's results. Consider stating the conclusion as "Computed tomography has substantially improved diagnostic performance over CXR in COVID-19. CT should be considered in the initial assessment for suspected COVID-19 instead of CXR if capacity allows and balanced against radiation exposure risk."

Many thanks, I have amended the abstract exactly as you suggest.

2. Table 3, the terms "Apparent prevalence" and "True prevalence" should be explained as to what prevalence the authors are referring to. This will improve clarity of the table.

Apparent prevalence refers to the predicted prevalence in the study by each imaging modality if they reported a Classical image, true prevalence refers to the number of positive cases in the study confirmed by RT-PCR. This has been added to the table caption.

3. The diagnosis of patients with COVID-19 is still unclear in the manuscript. Furthermore, for RT-PCR results to be valid for 30 days from initial ED attendance leaves many questions about the separation of the cohort into COVID-19 positive and negative. This needs to be listed as a limitation. It is well documented that COVID-19 patients can have negative RT-PCR results at presentation. If it is possible that some patients in this study could be in this category, this should also be stated as a limitation.

We have added this as a limitation, it may be the case that patients were not COVID-19 positive at ED attendance, but developed COVID-19 later. Where patients had negative RT-PCR at presentation they may have had a repeat RT-PCR within 30 days and subsequent positive results were taken as an overall positive, this has been described in the methods. It may be the case that some who tested negative initially may not have repeat tests however and this has been added as a limitation.

4. As the sensitivity of the RT-PCR test is unknown, please add references to the sensitivity of RT-PCR from previously published studies. Consider the following BMJ article to reference:

Jessica Watson, Penny F Whiting, John E Brush. Interpreting a covid-19 test result. BMJ 2020.

<https://www.bmj.com/content/bmj/369/bmj.m1808.full.pdf>

Many thanks we have added this.

5. As CXRs are all PA films, this should be stated in the methods section to allow the readers to understand that these are high quality CXRs unlike AP CXRs. Consider mentioning this in the discussion as this not the case in all centres worldwide where AP films are relatively common especially in unstable patients.

This has always been stated in the methods (line 3). I have also added this to the discussion.

6. To provide clarity of the CXR report gradings in table 2, it should state that this follows BSTI grading. Furthermore, the grades should be stated with the descriptions. For example, “Grade 0 – Alternative pathology”, “Grade 1 - No abnormalities”. In the results section a brief description of the number of patients with the gradings of the CXR and CT would provide further clarity to the paper.

Many thanks I have added these.

7. In Methods section under Diagnostic Accuracy Statistics (page 103 of 127 line 26), please include a reference to the BSTI COVID-19 CT reporting template.

This has been added.

8. Under the strengths and limitations in the discussion section (page 111 of 127, line 34), the following line “This study is the largest conducted on imaging in the COVID-19 pandemic” needs to be amended. Please see the following references of larger studies conducted on COVID-19 imaging to help amend the sentence:

- i) G Herpe, M Lederlin, et al. Efficacy of Chest CT for COVID-19 Pneumonia in France. Radiology (2020)
- ii) K De Smet, et al. Diagnostic Performance of Chest CT for SARS-CoV-2 Infection in Individuals with or without COVID-19 Symptoms. Radiology (2020)
- iii) Ng M, Wan E, Wong H et al (2020) Development and Validation of Risk Prediction Models for COVID-19 Positivity in a Hospital Setting. International Journal of Infectious Diseases (2020)

Many thanks, at the initial time of submission (prior to these studies’ publications) this study would have been the largest conducted to our knowledge. I have now amended the sentence to ‘one of the largest.’

9. Under the the strengths and limitations in the discussion section (page 112 of 127, lines 28-32), the comment of “This is likely the case worldwide however as many asymptomatic patients remain undiagnosed and those with mild disease are encouraged to self isolate rather than see medical professionals and be tested” should be adjusted as different localities have adopted mass screening programmes of asymptomatic patients at different timepoints including the UK but on a smaller scale. Examples, include Wuhan, China around May 2020 and a link below to the University of Nottingham asymptomatic screening programme.

<https://www.nottingham.ac.uk/news/university-of-nottingham-to-develop-asymptomatic-testing-service-to-help-control-hidden-spread-of-covid-19>

Many thanks, we have amended this to say ‘This is the case in the majority of localities worldwide’

10. Consider adjusting the statement in manuscript conclusion (page 114 of 127, line 10) to reflect the study findings and grammar “and should strongly be considered during the pandemic...” to “and should be considered ahead of chest x-ray in the initial assessment of COVID-19”

Many thanks we have replaced this exactly as you suggest.

VERSION 3 – REVIEW

| | |
|-------------------------|--|
| REVIEWER | Ming Yen Ng The University of Hong Kong Received funds from Bayer AG and Circle Cardiovascular Imaging |
| REVIEW RETURNED | 05-Oct-2020 |
| GENERAL COMMENTS | The authors have answered my queries to my satisfaction. I note that BMJ Open prefers just 2 revision cycles. But if I am allowed one minor comment, I would say that a sentence in the methods section on the definition of a positive or negative COVID-19 patients would help with the paper's clarity and for the reader to better understand the study. |

VERSION 3 – AUTHOR RESPONSE

We have clarified as per the reviewer's suggestion, the paragraphs now read:

RT-PCR of swabs were performed in laboratories either at our centre or at a public health laboratory (PHE Collindale, UK), according to published national standard operating procedures [13]. Subsequent RT-PCR swabs taken within 30 days of the initial swab in ED were also included. In the event of multiple RT-PCR swabs over 30 days, a single positive swab was taken as an overall positive test during one admission.

And:

The primary outcome is sensitivity and specificity of initial CXR, where it is reported as having classic COVID-19 features in the ED. This is compared with RT-PCR swabs (including subsequent swabs after initial attendance within 30 days as described above) as the reference standard for diagnosis of COVID-19.