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)	Cohort profile: The East London Health and Care Partnership Data Repository - Using novel integrated data to support commissioning and research
AUTHORS	Ronaldson, Amy; Chandakas, Evangelos; Kang, Qiongwen; Brennan, Katie; Akande, Aminat; Ebyarimpa, Irene; Wyllie, Eleanor; Howard, George; Fradgley, Richard; Freestone, Mark; Bhui, Kamaldeep

VERSION 1 - REVIEW

REVIEWER	paul roderick Faculty of Medicine University of Southampton Southampton UK
REVIEW RETURNED	16-Apr-2020

GENERAL COMMENTS	This paper describes a rich routine health care data set in two boroughs in East London, a total population of 700k which is young deprived and multi-ethnic. Data on acute care, primary care, community and mental health services and substance misuse have been linked. It discusses demographics, validation of prevalence measures. prevalence rates of mental and physical health conditions and illustrative example of factors associated with AE use and data completeness. It will be of significant interest to researchers, clinicians, commissioners and policy makers who are seeking to replicate such a dataset and or to use these data for research particularly to explore the frequency and impacts of combine mental and physical health problems. However to fully inform this diverse audience several areas about the process of establishing the data and details of the content could be expanded. Data governance. The authors say there was no public engagement. Why not? Is any planned? On what basis can the data be used for research purposes? What are the implications for GPDR? What agreements were made with NHS Digital? Substance misuse data are sensitive and I understood confidential and anonymous, how did they secure access and what governance issues were involved. . They plan to include sexual

	health, what are the issues for accessing and linking such data. . How did they address the issue of de-identification within LSOAs? How would researchers and others access the data? Comment on the timeliness of the data i.e. built in delays in receipt and cleaning, could they be reduced? What is the scale of population mobility, how handle this? Was there any local resistance to the data linkages? If so how was this overcome? Context There are a few examples of how the data have been used by local commissioners but more could be said on the policy context of reconfiguration of primary care into PCNs, local health care into Integrated Care Systems and the need for more population health analysis to support these changes. Are the authors aware of similar linked whole population data (eg SAIL in Wales).A strength is the population coverage and ability to capture all providers for the dined population. Data completeness How might they improve data completeness e.g for the population with no LTC, for BMI? Have others found similar BMI missingness? Future use Outline whether they intend to link to wider determinants of health such as housing, benefits, and air pollution. Have they reflected and shared the challenges of social care link up. Conditions. What guided the choice of 9 mental health and 15 physical health conditions? How did they establish T1 DM? What are liver and pancreatic complications?. Is any cancer diagnosis too broad? How did they handle the categories of low mood and depression, is there any overlap or progression? Is it worth cross check depression data against anti-depressant medication. Data Any data on duplication? How much of what items? Data completeness Ethnicity is 86% complete, comparison of those missing or not would be helpful. Suggest don't include with White group but as separate group. Alcohol use from primary care: are there other ways of collecting this from Read coding? Validation Suggest prevalence validation is not face validity. One wouldn't expect the prevalence to be the same as national rate given population demographics etc, as well as the methods used for the national studies, national comparison at least needs age sex standardisation or discussion of this if its not feasible. Need more detail of the PHE local data; how comparable are they in data definition and source? Are the variation true or methodological artefact? Data presentation Table 2

	Put in missing smoking and BMI and alcohol data? Future imputation is mentioned but not elaborated, this would be helpful and expand on Bayesian and imputation approaches. Have they built a core team of experienced statisticians? Don't put in mean BMI based on 24% Table 3 Confidence limits would be helpful for comparisons Reference sources. Table 4 AE attendance Analysis would benefit from statistics input, need confidence intervals on the rates. Why choose AE not admissions. Data are counts some with multiple, poisson method might be most appropriate, and with age sex adjustment.

REVIEWER	Wullianallur Raghupathi Fordham University, USA
REVIEW RETURNED	27-Apr-2020

GENERAL COMMENTS	The paper describes a rich new repository of health data. The paper can be sharply enhanced by re-writing the research question. Is it describing the repository and what it has; or, is the describing the use of it (demonstrating efficacy); or, does it use the repository to do some research. There is a lot of narrative - could you add 1-2 figures/tables (e.g. summarize available variables in a table). The paper does describe some of the data in the repository. However, there is no visualization - can you add a couple of charts of the demographics (desc. statistics) or some correlations? Lastly, conclusion can be sharper - link back to revised research question.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name

Paul Roderick
Institution and Country
Faculty of Medicine
University of Southampton
Southampton, UK

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

These have been declared on p1.

• This paper describes a rich routine health care data set in two boroughs in East London, a total population of 700k which is young deprived and multi-ethnic. Data on acute care, primary care, community and mental health services and substance misuse have been linked. It discusses demographics, validation of prevalence measures. prevalence rates of mental and physical health conditions and illustrative example of factors associated with AE use and data completeness.

• It will be of significant interest to researchers, clinicians, commissioners and policy makers who are seeking to replicate such a dataset and or to use these data for research particularly to explore the frequency and impacts of combine mental and physical health problems. However, to fully inform this diverse audience several areas about the process of establishing the data and details of the content could be expanded.

We thank the reviewer for these supportive comments and are keen to ensure that their concerns are addressed.

Data governance.

• The authors say there was no public engagement. Why not? Is any planned?

Thank you for drawing this to our attention. The following text was added at p8:

A clinical expert group was convened consisting of fifteen clinicians, commissioners and public health experts from primary care and secondary care (mental health) services. This group met on a quarterly basis throughout the initial profiling of the cohort and provided feedback into the analysis and emerging characteristics.

• On what basis can the data be used for research purposes? What are the implications for GDPR? What agreements were made with NHS Digital? Substance misuse data are sensitive, and I understood confidential and anonymous, how did they secure access and what governance issues were involved? They plan to include sexual health, what are the issues for accessing and linking such data? How did they address the issue of de-identification within LSOAs? How would researchers and others access the data?

The following text has been added, also at p8/9:

This dataset has been compiled for the health and care sector to use for system planning. It contains sensitive patient level data; therefore, it is important to ensure appropriate protections are in place. Currently plans are being made to deploy the data into an electronic platform for internal use: recognising the possible value this data holds for researchers, there may be scope to support analysis deemed to be in the public interest, with appropriate data protections in place. Currently, permissions must be sought on a case by case basis and approved by relevant IG governance groups.

• Comment on the timeliness of the data i.e. built in delays in receipt and cleaning, could they be reduced?

From discussions within the author team and the Tower Hamlets Clinical Commissioning Group it is our understanding that the process for gathering and collating data has been streamlined as much as is possible at the moment. We hope that Figure 1 outlines the complexity of the ‘data journey’. We are not entirely sure it is possible to get more detailed information relating to actual timelines for each leg of this journey.

• What is the scale of population mobility, how handle this?

We agree that trying to account for population mobility while trying to carry out longitudinal analyses on a dataset of this kind is important due to the nature of 'churn' within the primary care system. However, in this Cohort Profile we provided a cross-sectional snapshot so felt it was not necessary to discuss population mobility.

In order to assess population mobility within the dataset, the CCG have been using length of GP registration as a proxy for population mobility, i.e. a patient registered for less than a year at the time of analysis are 'new residents' whereas those registered for more than 10 years are considered to be 'long term residents'. They also combine some of this data with immigrant status. The main caveat with this is that it is not possible to determine whether this population mobility is within the borough only, or whether patients have come from other boroughs. This is a limitation of the dataset and we now mention it in the Strengths and Limitations section of the profile.

• Was there any local resistance to the data linkages? If so, how was this overcome?

Given the possible sensitive nature of this data there were naturally information governance (IG) concerns associated with linking it together. IG rules governing the data storage and use remain understandably strict. We worked with colleagues to address the IG concerns and the following texts have been added on p8:
"The dataset contains sensitive patient level data; therefore, it was and remains important to ensure appropriate protections are in place. All data are de-identified, with no information such as name, home address or NHS number. Also, certain sensitive health data such as records relating to trauma and abuse, or specific sexual health data, are not visible.
To further preserve anonymity and to ensure proportionate and appropriate use of the data, access to records are limited to only being able to view aggregated records where there is a minimum of 5 individuals. Searches where there are fewer than 5 records are suppressed. Access to the complete de-identified records is further limited to specific individuals who are employed by, or under contract with, the health and care sector, have knowledge of how to use the data effectively and have undergone appropriate training. "

Context
• There are a few examples of how the data have been used by local commissioners but more could be said on the policy context of reconfiguration of primary care into PCNs, local health care into Integrated Care Systems and the need for more population health analysis to support these changes. Are the authors aware of similar linked whole population data (eg SAIL in Wales)? A strength is the population coverage and the ability to capture all providers for the dined population.

Perhaps this is ignorance on behalf of the investigating team but we are not aware of these changes and none of the authors was able to offer a response.

• Data completeness
How might they improve data completeness e.g for the population with no LTC, for BMI? Have others found similar BMI missingness?

It appears that certain clinical markers are more frequently recorded for patients with LTCs who likely use primary care services more. This is largely down to primary care clinicians in practice and Improving data completeness for the population with no LTCs (e.g. for BMI) would have to come from clinicians themselves. This has been reported in the UK previously:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6377050/. Petersen and colleagues (2010) partially

attribute the discrepancy to GP incentive schemes. We now cite this paper in the Strengths and Limitations section of the paper. As we mention in the paper, it is possible to resolve this issue with multiple imputation to a certain degree.

• Future use
Outline whether they intend to link to wider determinants of health such as housing, benefits, and air pollution. Have they reflected and shared the challenges of social care link up.

A project to do this was trialed by the London Borough of Tower Hamlets as part of the NHS England designated 'Tower Hamlets Together Vanguard' program - established to support development of coordinated health and care systems and to capture learnings to share with others in the NHS. The datasets described above were successfully joined covering a limited time frame and on a trial basis. However, the combination of this data was deemed so sensitive that in practice only a handful of designated users were able to access the dataset and could only do so to answer a specific set of pre-determined questions. Given this, work has not been rolled out more widely.

• Conditions.
What guided the choice of 9 mental health and 15 physical health conditions?

The long-term condition variables available for analysis in the dataset were determined and created by the North East London Commissioning Support Unit (NEL CSU) using Read codes. NEL CSU were responsible for, in collaboration with Tower Hamlets CCG, developing the dataset. The justification for the choice of conditions included is information that would need to be obtained from NEL CSU. We can request this information if the reviewer would like.

• How did they establish T1 DM?

The T1DM variable was created by NEL CSU using C10 Read codes as detailed on p6-7. We are aware there are issues surrounding the use of this code in primary care (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC514025/) but this is an issue that would need to be rectified within clinical practice.

• What are liver and pancreatic complications?

This is another variable created by NEL CSU which comprises conditions such as liver cirrhosis, liver disease, chronic pancreatitis, and hepatitis. We agree this is quite an ambiguous 'catch-all' variable and have provided clarity in Table 1.

• Is any cancer diagnosis too broad?

We agree that use of cancer as a general term is broad and encompasses myriad diseases, and that this might be problematic in terms of fully understanding cancer using this dataset. This is a problem that seems to be quite common in studies using primary care data (https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(12)60240-2/fulltext#secd4414136e1139). Again, the cancer variable was created by NEL CSU using the relevant Read codes so there is little scope for change. However, they omitted basal cell carcinoma from the variable, likely because this is one of the least risky types of skin cancer.

We have now mentioned this in Table 1.

• How did they handle the categories of low mood and depression, is there any overlap or progression?

In the paper, we report figures of overall depression (which includes potential overlap with low mood), and figures of low mood excluding depression (see Figure 1). There is significant overlap between these two categories. We have now included an example of this from the Tower Hamlets data in the 'Prevalence rates of mental and physical health conditions' subsection of the Results.

• Is it worth cross check depression data against anti-depressant medication?

This is an excellent point and something we considered doing in order to capture those with depression who may not have a recorded diagnosis. However, after numerous consultations with primary care clinicians, it was decided that using antidepressant prescriptions to determine depression would be problematic seeing as antidepressants are often prescribed for conditions such as neuropathic pain and migraine. Certainly, antidepressant use within those with a depression (or low mood) diagnosis is doable, but we believe it is not within the scope of the Cohort Profile.

• Data
Any data on duplication? How much of what items?

Approximately 10% of the data captured in any snapshot was duplicates. The primary source of duplicate records was multiple assignations of patients to PBR clusters. To mitigate this all analysis was run with duplicates removed.

• Data completeness
Ethnicity is 86% complete, comparison of those missing or not would be helpful. Suggest don't include with White group but as separate group.

Unfortunately, the investigating team were unable to retain raw data after the initial cleaning and would not be able to produce a re-analysis of missing data+in a timely fashion. We agree that including 'not stated' with the White group is problematic but unfortunately this is the nature of how ethnicity is recorded in primary care systems in Tower Hamlets and City & Hackney. In 2011, 31% of Tower Hamlets residents where white (https://www.towerhamlets.gov.uk/Documents/Borough_statistics/Ward_profiles/Census-2011/RB-Census2011-Ethnicity-2013-01.pdf). In the current study (2017/2018), 38.2% were 'white or not stated' which is in line with government figures.

• Alcohol use from primary care: are there other ways of collecting this from Read coding?

As far as we aware (after liaising with clinical and commissioning colleagues) the AUDIT and AUDIT-C are the only widely used measures of alcohol use within primary care.

• Validation
Suggest prevalence validation is not face validity. One wouldn't expect the prevalence to be the same as national rate given population demographics etc, as well as the methods used for the national studies, national comparison at least needs age sex standardisation or discussion of this if its not feasible. Need more detail of the PHE local data; how comparable are they in data definition and source? Are the variation true or methodological artefact?

This is a good point, and we have now removed the phrase 'face validity' from the Methods section of the manuscript.

• Data presentation
Table 2 - Put in missing smoking and BMI and alcohol data?

To our knowledge smoking data is complete for all patients reported on in this cohort profile. Regarding BMI, would the reviewer like us to add a row indicating the % of missing data for BMI? Due to the considerable 'missingness' to do with alcohol use data (only reported in ~50% of patients without LTCs) we decided to omit alcohol use from Table 2 as we felt this would be misleading.

• Future imputation is mentioned but not elaborated, this would be helpful and expand on Bayesian and imputation approaches. Have they built a core team of experienced statisticians?

The study team are working with the KNIFE (KNowledge dIscovery From hEalth use Data - https://www.turing.ac.uk/research/research-projects/knowledge-discovery-health-use-data) project funded by the Alan Turing Institute which involves experts in AI, including one of the authors (Dr Mark Freestone). We have added this information on p14.

• Don't put in mean BMI based on 24%

BMI data available for ~70% of patients without LTCs and ~97% of those with LTCs. Therefore, we felt it was acceptable to report BMI in Table 2.

• Table 3
Confidence limits would be helpful for comparisons

We have added these to Table 3 as requested.

• Reference sources.

Thank you for spotting this – we have now added references to the sources.

• Table 4 AE attendance
Analysis would benefit from statistics input, need confidence intervals on the rates.

Confidence intervals for the crude rates have now been added to Table 5 (p. 22)

• Why choose AE not admissions.

A&E admissions were chosen as an example of the sort of health service utilisation data that is available. We have made it clearer on p.12 that other health service utilisation variables (e.g. inpatient admissions) data are available in the 'A&E attendances across mental health conditions' subsection of Findings to Date.

• Data are counts some with multiple, poisson method might be most appropriate, and with age sex adjustment.

We agree with the reviewer that this would be useful. However, the investigating team were unable to retain raw data after the initial cleaning and would not be able to produce this analysis in a timely fashion, particularly considering current nationwide coronavirus lockdown. Confidence intervals added to Tables 3 and 5 were calculated using standard formulae and did not require access to the raw data. We believe reporting raw counts and proportions is useful for showcasing the sort of data available within the linked data set nevertheless.

Reviewer: 2
Reviewer Name
Wullianallur Raghupathi
Institution and Country
Fordham University, USA
Please state any competing interests or state 'None declared':
None declared.
Please leave your comments for the authors below:

• The paper describes a rich new repository of health data. The paper can be sharply enhanced by re-writing the research question. Is it describing the repository and what it has; or, is the describing the use of it (demonstrating efficacy); or, does it use the repository to do some research.

We thank the reviewer for pointing out the ambiguity surrounding the research question. As this is a Cohort Profile there is not strictly a research question. Cohort profiles are used to describe and introduce novel datasets to inform research colleagues and to facilitate research and the creation of other similar datasets. In this paper we describe the dataset, how it was formed, what sort of information is included in its current form, and what the current state of the data is (i.e. representativeness and completeness). We understand that the purpose of the Cohort Profile may not have been clear and have therefore added some further information to make this clear on p5.

• There is a lot of narrative - could you add 1-2 figures/tables (e.g. summarize available variables in a table).

This Cohort Profile comprises six tables and four figures in its existing form which is toward the upper limits in terms of publication norms. We believe once these tables and figures are placed throughout the published paper in a way that makes sense this will break up the narrative and make the information more digestible.

• The paper does describe some of the data in the repository. However, there is no visualization - can you add a couple of charts of the demographics (desc. statistics) or some correlations?

There are several tables and figures which describe demographics and prevalence rates (Tables 2, 3 and 4, Figures 2, 3, and 4). If the reviewer would like us to reformat the way in which we present the information we would be happy to do so. However would need some editorial guidance on how best to do this.

• Lastly, conclusion can be sharper - link back to revised research question.

We have now added a sentence to the Conclusion paragraph to make the purpose of the paper a little clearer.

VERSION 2 – REVIEW

REVIEWER	Paul Roderick University of Southampton
REVIEW RETURNED	09-Jul-2020

GENERAL COMMENTS	The authors have responded to the previous referees' comments There are still a few issues raised that could be further clarified i) How do researchers who are interested in using the data apply?

	ii) Previous referee comment: There are a few examples of how the data have been used by local commissioners but more could be said on the policy context of reconfiguration of primary care into PCNs, local health care into Integrated Care Systems and the need for more population health analysis to support these changes. Are the authors aware of similar linked whole population data (eg SAIL in Wales)? A strength is the population coverage and the ability to capture all providers for the dined population. Authors' Reply : Perhaps this is ignorance on behalf of the investigating team but we are not aware of these changes and none of the authors was able to offer a response. This seems weak given the collaboration with the CCG/commissioners, could this be pursued iii) Previous referee comment Future use Outline whether they intend to link to wider determinants of health such as housing, benefits, and air pollution. Have they reflected and shared the challenges of social care link up. Authors' reply A project to do this was trialed by the London Borough of Tower Hamlets as part of the NHS England designated 'Tower Hamlets Together Vanguard' program - established to support development of coordinated health and care systems and to capture learnings to share with others in the NHS. The datasets described above were successfully joined covering a limited time frame and on a trial basis. However, the combination of this data was deemed so sensitive that in practice only a handful of designated users were able to access the dataset and could only do so to answer a specific set of pre-determined questions. Given this, work has not been rolled out more widely. Suggest mention this briefly as its an important finding. iv) Previous referee comment Validation Suggest prevalence validation is not face validity. One wouldn't expect the prevalence to be the same as national rate given population demographics etc, as well as the methods used for the national studies, national comparison at least needs age sex standardisation or discussion of this if its not feasible. Need more detail of the PHE local data; how comparable are they in data definition and source? Are the variation true or methodological artefact? Authors reply. This is a good point, and we have now removed the phrase 'face validity' from the Methods section of the manuscript. This could be expanded. Validity is still mentioned in the abstract. Need at least to discuss no age sex standardisation and the comparability of the local and national data. Assessing representativeness is difficult especially a such as young, socially deprived and ethnically diverse population. A specific point the high asthma prevalence is suggested as an outlier and artefact , could it be true finding due to high levels of air

	pollution?

REVIEWER	Wullianallur Raghupathi Fordham University, USA
REVIEW RETURNED	06-Jul-2020

GENERAL COMMENTS	The authors have met the suggested changes.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name
Paul Roderick

Institution and Country
University of Southampton

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

Please leave your comments for the authors below:

The authors have responded to the previous referees' comments. There are still a few issues raised that could be further clarified:

i) How do researchers who are interested in using the data apply?

Contact details have now been provided on both the Title Page and in the 'Conclusions and future collaborations' section of the Discussion.

ii) Previous referee comment: There are a few examples of how the data have been used by local commissioners but more could be said on the policy context of reconfiguration of primary care into PCNs, local health care into Integrated Care Systems and the need for more population health analysis to support these changes. Are the authors aware of similar linked whole population data (eg SAIL in Wales)? A strength is the population coverage and the ability to capture all providers for the dined population. Authors' Reply: Perhaps this is ignorance on behalf of the investigating team but we are not aware of these changes and none of the authors was able to offer a response.
This seems weak given the collaboration with the CCG/commissioners, could this be pursued

We have used this Cohort Profile to illustrate what can be done with linked datasets at CCG level, thus informing thinking about how integrated care services might be organised, and signposting where integrated care might be helpful (e.g. we have a forthcoming paper on diabetes, depression, and emergency department use). However, the data in its current form differs to that of SAIL and cannot be used to develop predictive models.

iii) Previous referee comment
Future use

Outline whether they intend to link to wider determinants of health such as housing, benefits, and air pollution. Have they reflected and shared the challenges of social care link up.

Authors' reply A project to do this was trialed by the London Borough of Tower Hamlets as part of the NHS England designated 'Tower Hamlets Together Vanguard' program - established to support development of coordinated health and care systems and to capture learnings to share with others in the NHS. The datasets described above were successfully joined covering a limited time frame and on a trial basis. However, the combination of this data was deemed so sensitive that in practice only a handful of designated users were able to access the dataset and could only do so to answer a specific set of pre-determined questions. Given this, work has not been rolled out more widely.

Suggest mention this briefly as it's an important finding.

We agree that it is important to include information about integration of sensitive data. We have now added a line to the 'The development of the ELHCP Data Repository' section (p. 5) stating the following: 'However, challenges surrounding linking social care data means that sensitive information relating to housing and benefits will not be integrated into the dataset'.

iv) Previous referee comment Validation
Suggest prevalence validation is not face validity. One wouldn't expect the prevalence to be the same as national rate given population demographics etc, as well as the methods used for the national studies, national comparison at least needs age sex standardisation or discussion of this if its not feasible. Need more detail of the PHE local data; how comparable are they in data definition and source? Are the variation true or methodological artefact?

Authors reply. This is a good point, and we have now removed the phrase 'face validity' from the Methods section of the manuscript.

This could be expanded. Validity is still mentioned in the abstract. Need at least to discuss no age sex standardisation and the comparability of the local and national data.
Assessing representativeness is difficult especially a such as young, socially deprived and ethnically diverse population.

We have now removed the reference to validity from the Abstract. We agree that assessing representativeness of data is difficult with the sort of unique cohort we describe. However, we compared the cohort prevalence data with PHE data from the same local area meaning that comparison can be made.

We agree that in an ideal situation prevalence data would be both age- and sex-standardised. However, the PHE prevalence data is not standardised. Therefore, we chose to present raw prevalence rates in the Cohort Profile so that direct comparison could be made (Table 3). We have now added a line to explain why we have not provided age- or sex- standardised prevalence data: 'We did not provide age- and sex-standardised prevalence rates from the ELHCP Data Repository to allow direct comparison with unstandardized local prevalence data from PHE' (p.10)

A specific point the high asthma prevalence is suggested as an outlier and artefact , could it be true finding due to high levels of air pollution?

This is an interesting point and it could be an explanatory factor in the high prevalence rates of asthma we have described in the ELHCP Data Repository. However, if air pollution was a driving factor one would think that this would also be reflected in the PHE data for the local area. PHE prevalence rates are taken directly from QOF registers which might explain the discrepancy. We have now added the following to p.10/11: 'Moreover, PHE prevalence data is taken directly from Quality Outcomes Framework (QOF) registers. Prevalence data from QOF registers might differ from data from other sources due to coding or definitional issues, as well as population churn.'

Reviewer: 2
Reviewer Name
Wullianallur Raghupathi
Institution and Country
Fordham University, USA
Please state any competing interests or state 'None declared':
None declared.
Please leave your comments for the authors below
The authors have met the suggested changes.

VERSION 3 – REVIEW

REVIEWER	paul roderick Faculty of Medicine University of Southampton Southampton UK
REVIEW RETURNED	06-Aug-2020

GENERAL COMMENTS	The authors have responded well to the feedback.