PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Research Evaluation Alongside Clinical Treatment in COVID 19 (REACT COVID 19): An Observational and Biobanking Study
AUTHORS	Burke, Hannah; Freeman, Anna; Dushianthan, Ahilanandan; Celinski, Michael; Batchelor, James; Phan, Hang; Borca, Florina; Kipps, Christopher; Thomas, Gareth; Faust, Saul; Sheard, Natasha; Williams, Sarah; Fitzpatrick, Paul; Landers, Dónal; Wilkinson, Tom

VERSION 1 – REVIEW

REVIEWER	Sergio Bonini Italian National Research Council
REVIEW RETURNED	12-Sep-2020

GENERAL COMMENTS	The paper describes the protocol of an observational prospective and biobanking study of patients admitted for COVID-19 at the University Hospital Southampton NHS Foundation Trust (the REACT COVID 19 study). Since only subjects tested positive for SARS CoV-2 who require hospitalization are included in the study, information on the clinical features and outcome of the large proportion of subjects infected by SARS- CoV-2 with no or minor clinical symptoms will not be provided by this study design. The objectives of REACT COVID-19 are clearly defined, however they are quite ambitious to be answered on the basis of data collected in the necessarily limited number of patients referred to a single institution. Multicenter international studies with common operational procedures and using the same platform might better answer the research question of REACT COVID- 19. There is no mention of the policy of authors and of the UHSFT about data sharing. This information is requested by all major journals and appears to be extremely important to make data available from the study of interest for the scientific co

REVIEWER	So, Hon-Cheong The Chinese University of Hong Kong
REVIEW RETURNED	18-Nov-2020

GENERAL COMMENTS	Comments to authors: This is a study protocol for an observational and biobanking study in the UK for COVID-19. This is a timely study and the study plan is well-described. The protocol is well-written and I do not find any major problems. Some suggestions are as follows:

	1) Introduction part: I would advise to update the figures listed in the 1st paragraph. 2) It was mentioned that samples stored in the biorepository (for example blood, urine, sputum) can be analyzed to develop an endotype level understanding of disease clusters. This is a promising direction to find subtypes of the disease; on the other hand, one may also consider using them for predicting disease outcomes. 3) For the study sample, I do not have major comments. Could the authors briefly comment on how representative the sample will be? For example, is there any risk of selection bias and how well can the findings be generalized to other UK population (or other populations)? 4) The subjects will be FU for 12 months following discharge. May I know if there are plans to FU longer, eg some adverse effects may last longer especially for those more severely ill? 5) For the statistical analysis, just a few suggestions a) I think authors may also mention time-to-event analysis (eg time to developing a certain adverse outcome) and longitudinal analysis (eg mixed models/GEE models) on some longitudinally measured outcomes (ie to study the trajectory of outcomes) b) Optionally, the authors may consider supervised learning analysis (ie prediction of outcome) in addition to clustering (unsupervised learning) c) For “Bayesian analysis part”, eg p.17 “With a possible extension of the Bayesian framework into a causal inference setting” “Provide a critical analysis of the ethical and safety aspects related to the clinical application of Bayesian inference, i.e. supporting the clinical decision-making process” Could you explain further what kind of Bayesian analysis were planned? Or provide some relevant references?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Reviewer: The paper describes the protocol of an observational prospective and biobanking study of patients admitted for COVID-19 at the University Hospital Southampton NHS Foundation Trust (the REACT COVID 19 study). Since only subjects tested positive for SARS CoV-2 who require hospitalization are included in the study, information on the clinical features and outcome of the large proportion of subjects infected by SARS-CoV-2 with no or minor clinical symptoms will not be provided by this study design:

Author: We thank the reviewer for their comment. The authors appreciate that this is a limitation to the study design, but also acknowledge that the population requiring hospital admission are those who are most unwell and likely to contribute most of the healthcare costs during the pandemic.

Therefore, this is an important group of patients to study in their own right but with the awareness of other presentations of SARS-CoV-2 infection.

Reviewer: The objectives of REACT COVID-19 are clearly defined, however they are quite ambitious to be answered on the basis of data collected in the necessarily limited number of patients referred to a single institution. Multicenter international studies with common operational procedures and using the same platform might better answer the research question of REACT COVID- 19.

Author: We thank the reviewer for this comment and acknowledge that this is a consideration. There is potential to expand this study to a multicentre study, but validation of its feasibility and utility on one site is the aim in the first instance. There is also clinical utility for UHS to be able to interrogate their patient cohort in granular detail in isolation.

Reviewer: There is no mention of the policy of authors and of the UHSFT about data sharing. This information is requested by all major journals and appears to be extremely important to make data available from the study of interest for the scientific co

Author: Apologies for this oversight. This has now been amended in the manuscript

Reviewer 2

Reviewer: This is a study protocol for an observational and biobanking study in the UK for COVID-19. This is a timely study and the study plan is well-described. The protocol is well-written and I do not find any major problems. Some suggestions are as follows:

Author: We thank the reviewer for their comments which we have now amended within the manuscript and responded to point by point below.

Reviewer: 1) Introduction part: I would advise to update the figures listed in the 1st paragraph.

Author: We thank the reviewer for this comment which we have now amended within the manuscript.

Reviewer: 2) It was mentioned that samples stored in the biorepository (for example blood, urine, sputum) can be analyzed to develop an endotype level understanding of disease clusters. This is a promising direction to find subtypes of the disease; on the other hand, one may also consider using them for predicting disease outcomes.

Author: Many thanks for the suggestion, the authors agree this is a potential additional use and have amended the manuscript accordingly.

Reviewer: 3) For the study sample, I do not have major comments. Could the authors briefly comment on how representative the sample will be? For example, is there any risk of selection bias and how well can the findings be generalized to other UK population (or other populations)?
Author: Many thanks for the comment. There is a risk of selection bias in that the study will only capture data on those patients unwell enough to present to hospital. However, this is a group towards which most treatment strategies will be targeted, and will carry the highest morbidity and mortality, and associated healthcare costs. The group may not be entirely generalisable to the UK population due to differering demographic distribitions across the UK, but there is potential to extend to a multicentre study with appropriate approvals.

Reviewer: 4) The subjects will be FU for 12 months following discharge. May I know if there are plans to FU longer, eg some adverse effects may last longer especially for those more severely ill?
Author: This was considered by the authors but the view of the REC was that consent would be needed to follow up beyond 12 months.

For the statistical analysis, just a few suggestions

Reviewer: a) I think authors may also mention time-to-event analysis (eg time to developing a certain adverse outcome) and longitudinal analysis (eg mixed models/GEE models) on some longitudinally measured outcomes (ie to study the trajectory of outcomes)
Author: Yes, we agree that we can certainly investigate time-to-event and longitudinal analysis of the patient outcomes in this study as we have these data captured in the UHS dataset.

Reviewer: b) Optionally, the authors may consider supervised learning analysis (ie prediction of outcome) in addition to clustering (unsupervised learning)
Author: Yes, we can explore additional Machine Learning and AI methods as part of the follow-up analysis, with a particular emphasis on explainable methods for clinical analysis (as defined by Lundberg et al., Nature Intelligence, 2020).

Reviewer: c) For "Bayesian analysis part", eg p.17 "With a possible extension of the Bayesian framework into a causal inference setting"

"Provide a critical analysis of the ethical and safety aspects related to the clinical application of Bayesian inference, i.e. supporting the clinical decision-making process"

Could you explain further what kind of Bayesian analysis were planned? Or provide some relevant references?
Author: Our main aim was to cover a broader methodological ground beyond uni/multi-variable statistical analysis and the direct application of black-box ML models, comparing and contrasting their interpretability and safety aspects within the clinical problem space. Two methods which point into the causal inference space were planned to be instantiated within this dataset: the causal framework of Pearl (Bayesian-based) and the information-theoretical model for feature selection proposed in Sechidis et al. 2019. Our main goal with the first framework was to elicit the modelling/experimental conditions required for a causal claim in this problem space (we do not expect that the dataset will enable these claims to be stated). As noted, we understand that the Bayesian frameworks are sensitive to the encoding and assumptions of priors, which need to be grounded either on solid scientific evidence or on most conservative estimates. Our main motivation with the Bayesian modelling was to support a more explicit modelling of the space of random variables and their dependencies associated with the dataset.

VERSION 2 – REVIEW

REVIEWER	Hon-Cheong So The Chinese University of Hong Kong, HK , China
REVIEW RETURNED	28-Dec-2020

GENERAL COMMENTS	My concerns are addressed adequately. Thank you.