COVID RESPONSE STUDY (COVRES)

Research Protocol

NCT XXX

13/08/21

COVID RESPONSE STUDY (COVRES)

A Northern Ireland population study of SARS-CoV-2 prevalence, predisposing factors and pathology

Project Team

Professor Tony Bjourson, Director of the Northern Ireland Centre for Personalised Medicine, Ulster University (CI)

Dr David Gibson, Lecturer in Stratified Medicine, School of Biomedical Sciences, Ulster University

Professor Colum Walsh, Professor of Genetics, School of Biomedical sciences, Ulster University

Professor Stuart Elborn, Faculty Pro-Vice-Chancellor, School of Medicine, Dentistry and Biomedical Sciences, Queens University

Dr Magda Bucholc, Lecturer in Data Analytics, School of Computing & Eng & Intel. Sys, Ulster University

Dr Shu-Dong Zhang, Senior Lecturer School of Biomedical Sciences, Ulster University

Dr Taranjit Singh Rai, Lecturer in Cellular Ageing, School of Biomedical Sciences, Ulster University

Dr Vanessa Devine, GMI Programme Manager, Western Health and Social Care Trust

Principal Investigators

Dr Maurice O'Kane, Consultant Chemical Pathologist, Western Health and Social Care trust

Dr Peter Sharpe, Consultant Chemical Pathologist, Southern Health and Social Care Trust

Dr Derek Fairley, Research Scientist, Belfast Health and Social Care Trust

Background

The SARS-CoV-2 pandemic has caused greater than 5 million worldwide infections up to 4th June 2020. In Northern Ireland (NI) as of 4th June 2020, 55,909 people have been tested by PCR assay for presence of the virus and 4,773 positive cases have been identified and 535 deaths. However, as the UK testing strategy remains concentrated on hospitalised cases and healthcare workers and their families, community spread of the infection and possible deaths remains unknown. NI positive cases by age group 2.3% in 0-19; 22.8% in 20-49; 19.7% in 60-79 and 25.6% in the over 80 age group (data from http://nicovidtracker.org/ 11/05/20) and gender (60% Female, 40.0% male) indicate that predisposing factors may be

at play in determining risk of infection. While more females test positive, males are twice as likely to die from SARS-CoV-2 compared to females. Additionally, several studies suggest that variable immune and inflammatory responses during infection may play a role in likelihood of developing pneumonia and subsequent recovery.

There is therefore an urgent need to more clearly understand: (a) the prevalence of the infection in the wider community in Northern Ireland (b) genetic and comorbidity factors predisposing individuals to more severe infection and higher risk of death (c) immune and inflammatory response during active infection and recovery.

With a clearer understanding of these SARS-CoV-2 characteristics at both population and biology levels we will be better equipped to cater our ongoing regional response. This study could help inform the regional strategy of exit from lock-down, identify at risk individuals with specific predisposing factors at a global level and advance our knowledge of the underlying genetics and immune biology which dictate response to infection. It is hoped that understanding the pathology of SARS-CoV-2 could help to reduce the mortality rate amongst individuals who are particularly vulnerable to this infection. This research is a collaboration between Ulster University and Queens University and will take place across three trusts areas in N. Ireland namely Western Health and Social Care Trust (WHSCT) to include The South West Acute Hospital, Fermanagh, Southern Health and Social Care Trust (SHSCT) and Belfast Health and Social Care Trust (BHSCT).

Overall Purpose/objectives

The overall purpose of the COVRES study is to provide preliminary data on prevalence and factors influencing severity of the infection in the NI population.

The core objectives are:

a. To measure IgG to SARS-CoV-2 in a representative sample of the NI population at community level (nursing home), in hospitalised and non-hospitalised SARS-CoV-2 positive cases and in frontline healthcare staff.

b. To perform whole genome sequencing on phlebotomy blood samples from hospitalised SARS-CoV-2 positive cases and frontline healthcare staff and collect associated SARS-CoV-2 PCR test data, comorbidity and medication data.

c. To analyse immune cell and inflammatory responses in hospitalised and non-hospitalised SARS-CoV-2 positive cases and frontline healthcare staff.

d. explore the relationship between host genotype, viral genotype, phenotypic and other data with (i) predisposition to Covid19 infection, (ii) severity (iii) treatment response if appropriate: drug/Vaccine/intervention and (iv) disease epidemiology.

Study Design: This study will undertake: (a) the collection of retrospective waste plasma/serum samples from Covid-19 positive patients and extract from existing nasopharyngeal swabs and (b) the recruitment of Covid-19 positive hospitalised and non-

hospitalised and nursing home patients, hospital and nursing home staff, for whole genome sequencing, human host genomic analyses and viral genomic analyses. This research study will take place across three trust areas in Northern Ireland and will be conducted simultaneously in four phases, namely:

Phase 1- Rapid access of waste serum/plasma samples from hospital laboratories
Phase 2- Rapid access of nasopharyngeal swab extract, collected from patients, NHS and nursing home staff and members of the public.
Phase 3- Recruitment of SARS-CoV-2 patients who have tested positive by virus PCR assay from both hospitals and trust affiliated nursing homes and community.
Phase 4- Recruitment of hospital-tested and nursing home staff

Recruitment of participants and sample collection

Rapid Access to retrospective samples from hospital laboratories:
-Waste serum/plasma samples (routine patient samples that have been analysed and clinically acted upon and awaiting destruction) will be retained and stored by the laboratory staff and on a weekly basis collected by the researcher(s). The samples will be used in the validation of antibody assays. In collaboration with the Biomedical Scientist at each laboratory site approximately 2ml of anonymised sample per patient will be retained for the research team, details required initially would be time and date of sample collection, age, gender and whether the sample tested PCR positive or negative. Guided by MHRA requirements to validate antibody assays it would be expected that n = 200 PCR negative and n = 200 PCR positive would be required. In addition, we will use n = 200 pre-pandemic control samples already retained from previous genetic studies at the Northern Ireland Centre for Stratified Medicine (NICSM), C-TRIC building, Altnagelvin Hospital site, Londonderry and for whom the participants had given enduring consent for their samples to be used in other genetic research. As these samples will be collected retrospectively, it would be imperative to follow up with these participants at three, six and 12 months to ask them to provide a saliva sample, 30ml plasma sample but specifically to check if in those who tested positive a) The antibody persists and b) to extract DNA (which is not available in waste sample) and provide a saliva sample (~4ml). In addition, one set of five 20ul dried blood spots will also be collected at recruitment at each timepoint. NICRN nurse/research nurse will liaise with the PI and the clinical care team(s) based at each trust to initially follow up with those patients whose samples were used in the validation process: For those who are contactable (still an inpatient, recovered and discharged from hospital or tested negative) the nurse will either call to the potential participant on the ward or telephone them at home to explain the study, what would be required and gain verbal consent for the researcher to send them out a PIS (appendix 2). For those who choose to participate or hear more about the research, they can then contact the researcher whose details are on the PIS.

Informed written consent (appendix 3) will be obtained prior to enrolment on study. In the case of Adults Lacking in Capacity (ALC), and where a family member is not available to act as a nominated person the research team in conjunction with the PI will identify a nominated person who is unrelated to the conduct of the research and willing to provide a declaration for this study. The nominated person will always be a member of the nursing/medical staff or the management team and part of the wider multi-disciplinary team within the hospital ward or nursing home facility of the individual lacking capacity. They (the nominated person) will also be provided with an invitation letter and PIS (appendix 11) and any queries answered, the nominated person can then contact the research team if they think the potential participant would have been interested in this research. The potential participant will not be enrolled on the study until full informed consent (appendix 12) has been given by the nominated person.

Rapid Access to retrospective samples from nasopharyngeal swabs

-Rapid access will be also be sought across the three trusts for extract from the nasopharyngeal swabs collected from patients, NHS and nursing home staff and members of the public throughout the three Trust areas (these swabs will have already tested COVID-19 positive or negative and results acted upon).The extract will undergo viral genome analysis. 16S microbiome analysis will also be carried out on a selected COVID positive subset with matching controls to correlate microbiome with severity. Up to 500 sample whole genome sequencing and other genomic analyses will be performed. Again, as these samples will be collected retrospectively, it would be imperative to follow up with a subset of these participants (~ n = 500) at three, six and 12 months and ask them to provide a saliva sample, 30ml plasma sample to check if a) The antibody persists and b) to extract DNA (which is not readily available in the swab extract) and provide a saliva sample (~4ml). In addition, one set of five 20ul dried blood spots will also be collected and a lateral flow assay for COVID-19 antibodies performed at recruitment and at each timepoint. NICRN/research nurses based at each trust will initially liaise with the custodians of the extract to follow up with those patients, for those who are contactable (recovered and discharged from hospital, didn't need hospitalisation or tested negative) the nurses will post out a, pre-packed, PIS and invitation letter (appendix 4) about this research, a consent form (appendix 5) and a stamped addressed return envelope. Those that are interested can call the researchers to find out more and if still interested post back the consent form in the pre addressed envelope, the researchers will then follow up. For those patients who are hospitalised or in a nursing home a member of the clinical care team can discuss the study with the patient and if they are interested, get verbal permission for the researchers to give them a PIS (appendix 4). For those who wish to participate, they can contact the researcher whose details are on the PIS. Informed written consent (appendix 5) will be obtained prior to enrolment on study. In the case of those ALC, and where a family member is not available to act as a nominated responsible person, the research team in conjunction with the PI will

identify a nominated responsible person who is unrelated to the conduct of the research and willing to provide a declaration for this study. The nominated person will always be a member of the nursing/medical staff or the management team and part of the wider multi-disciplinary team within the hospital ward or nursing home facility of the individual lacking capacity. They (the nominated responsible person) will also be provided with an invitation letter and PIS (appendix 11) and any queries answered, the nominated person can then contact the research team if they think the potential participant would have been interested in this research. The potential participant will not be enrolled on the study until full informed consent (appendix 12) has been given by the nominated person. This will still apply at all time points for those individuals who are unlikely to regain capacity and for those patients in hospital who are sedated/ventilated due to their COVID-19 symptoms, who may regain capacity. For those patients, capacity will always be assessed by a member of their direct clinical care team, the nurse in charge/manager in the nursing home and the PI at each hospital site who will be the consultant on or linked to the COVID ward **(The person assessing capacity will be recorded at all time points and a record kept).** For those that regain capacity during the duration of this research they will be given participant information sheet (appendix 4) to read for themselves and their consent (appendix 5) will be sought if they wish to proceed with the study. If they do not wish to continue, they will be given the option for any samples and data to be withdrawn and destroyed.

Hospital and nursing home patient recruitment:

COVID-19 positive/ SARS-CoV-2 inpatients and non-hospitalised patients (n = 700) aged 18 years and over tested at the WHSCT, BHSCT and SHSCT and have been admitted to COVID-19/ SARS-CoV-2 hospital wards (including from the Intensive Care Units) or out-patient and residents of trust affiliated nursing homes will initially be approached by their clinician, senior nurse or an NICRN research nurse. They will speak to the patient(s) about the study and if the patient is interested, verbal consent will be gained for the researcher to follow up with a Participant Information sheet (PIS) (appendix 6). Those interested patients fulfilling the inclusion criteria can then speak directly face to face (whereby the nurse will arrange a convenient time for researcher to visit the patient) to the researchers, the study will be discussed in detail and any issues will be addressed. If they are still interested in participating, written informed consent (appendix 7) will be obtained from the patient by the researcher(s) prior to enrolment on the study.

In the case of those ALC, and where a family member is not available to act as a nominated person the research team in conjunction with the PI will identify a nominated person who is unrelated to the conduct of the research and willing to provide a declaration for this study. The nominated person will always be a member of the nursing staff or the management team and part of the wider multi-disciplinary team within the hospital ward or nursing home facility of the individual lacking capacity. They (the nominated person) will also be provided

with an invitation letter and PIS (appendix 13) and any queries answered, the nominated person can then contact the research team if they think the potential participant would have been interested in this research. The potential participant will not be enrolled on the study until full informed consent (appendix 14) has been given by the nominated person. This will still apply at all time points for those individuals who are unlikely to regain capacity and for those patients in hospital who are sedated/ventilated due to their COVID-19 symptoms, who may regain capacity. For those patients, capacity will always be assessed by a member of their direct clinical care team, the nurse in charge/manager in the nursing home and the PI at each hospital site who will be the consultant on or linked to the COVID ward **(The person assessing capacity will be recorded at all time points and a record kept).** For those that regain capacity during the duration of this research they will be given participant information sheet (appendix 4) to read for themselves and their consent (appendix 5) will be sought if they wish to proceed with the study. If they do not wish to continue, they will be given the option for any samples and data to be withdrawn and destroyed.

Sample Collection: Once informed consent has been obtained each participant will provide (i) one 30 ml whole blood sample, a saliva sample (~4ml) and one set of up to five 20ul dried blood spots (DBS; equivalent to five finger lancet blood droplets) and a lateral flow assay for COVID-19 antibodies performed at recruitment (between 0 and 10 days of symptom onset), (ii) again between 14-28 days after symptom onset (iii) again samples will be collected only if pneumonia occurs during episode (usually only in severe symptom cases) (iv) In addition, where possible samples will be collected at post 3 months, 6 months, 12 months and 18 months post symptom onset, the latter sample collections will take place at an agreed convenient time and place.

Hospital and nursing home staff recruitment: All HSC employees aged 18 years and over, employed within facilities at the WHSCT, BHSCT and the SHSCT and trust affiliated nursing homes within the trust areas (n = 500) are eligible to participate in this research. Trust/nursing home wide emails with an attached PIS explaining the study (appendix 6) will be sent to all staff. In addition, information regarding the research will be put on Trust and Nursing home websites and social media platforms e.g. Facebook and Twitter. Posters (appendix 8) will also be placed in communal areas ensuring information can be accessed by everyone. Staff will be directed to the researchers on the study via email or phone to discuss the study in more detail. For those who still wish to participate and who meet the inclusion criteria, written informed consent (appendix 7) will be obtained by the researcher(s).

Sample Collection: Once informed consent has been obtained each participant will provide (i) one 30 ml whole blood sample, one set of dried blood spots (DBS) and a saliva sample (~4ml) at recruitment. Then, only if SARS-CoV-2 symptoms such as a high temperature/a

new continuous cough https://www.nhs.uk/conditions/coronavirus-covid-19/check-if-you-have-coronavirus-symptoms/ experienced during 2 week post recruitment period: (ii) a saliva sample, one 30 ml whole blood sample and one set of dried blood spots (DBS) and a lateral flow assay for COVID-19 antibodies performed (between 0 and 10 days of symptom onset), (iii) a saliva sample, one 30 ml whole blood sample and one set of DBS and a lateral flow assay for COVID-19 antibodies performed between 14-28 days after symptom onset (iii) a saliva sample, one extra set of 30ml whole blood and DBS to be taken only if pneumonia occurs during (usually only in severe symptom cases) (iv) In addition, where possible sample will also be collected at 3 months, 6 months and 12 months of symptom onset. Sample collections will take place at an agreed convenient time and place.

Data Collection for all cohorts:

The phenomic data collected will align with international/National data for phenotypic data collection (HDRUK). After informed consent (appendices 3, 5 & 7) is obtained by a member of the research team, participants will be asked to answer a health and lifestyle questionnaire at each timepoint (Appendix 9) that includes questions about their general medical and clinical history such as comorbidities, medications, respiratory illness, COVID-19 symptomatology, time to admission to hospital (if hospitalised) as well as lifestyle habits, family history etc.

-Missing information will be collected by the researchers by accessing the participants electronic care records or the participants medical notes provided they have consented for this option. Demographic variables in addition to age and gender will include height, weight and blood pressure will also be retrieved from the participants electronic care record. The researchers will have an honorary contract within the Health Trust(s) and will have approval and training to access and to use the patient care records. In case certain information is missing, the researchers will ask permission to follow up with the patient at a later time by phone. All medical records will be accessed on the Altnagelvin hospital site at C-TRIC and information will be manually inputted to a database on a university password protected and encrypted laptop.

Participants will also be asked whether they consent at each time point, to complete a validated mental health questionnaire, the General Health Questionnaire (GHQ-12), (appendix 10) to assess the association between their condition or perceived risk and their mental health and wellbeing. The researchers will arrange for completion of this questionnaire at a convenient time for the participant either on the day of recruitment, during the follow up visits, online or by phone.

-Participants also have the choice to consent to be contacted again to receive information regarding future studies.

Sample processing, storage and analysis

Saliva ~ 4ml will be collected in Oragene saliva collection kits (DNA Genotek), processed and analysed for COVID-19 genome sequencing and microbiome analysis.

Whole blood samples will be processed to extract plasma and flow cytometry. Paper discs will be punched out of dried blood spot samples (DBS) and protein, prescribed medication, drug metabolites and DNA extracted by addition of a routine elution buffer. Lateral flow data will be logged at the time at of assay and where possible dedicated COVID-19 app photograph of test strip result archived. DBS extracts and plasma samples will be logged, aliquoted and labelled with unique bar code identifiers prior to storage at -80'C in the Northern Ireland Centre for Stratified Medicine (NICSM) until analysis. DBS protein (IgG, C-reactive protein and cytokines and plasma will be analysed using ELISA and OLINK or drug metabolites by mass spectrometry. With participant agreement, DNA extracted from DBS/blood samples for future whole genome sequencing (WGS) and extracted drug metabolites will be stored at -80^{0}C in. A whole genome sequence will be generated from each participant DNA sample. A selected cohort of patient samples will also be subjected to transcriptomics (RNA Seq), microbiome and methylation analyses.
Whole blood samples and swab extract will undergo viral genome analyses sequencing, transcriptomics, epigenetics, single nucleotide polymorphism (SNP) genotyping, WGS and RNAseq generation**.** Sample analysis will be performed at NICSM by the researchers or by a third party such as the virology laboratories at Queens University Belfast (QUB), Regional Virus Laboratory Belfast (RVL), Olink® proteomics (Uppsala, Sweden) or Genomics Medicine Ireland (GMI, Dublin, Dublin). This will require aliquots of plasma/serum or DNA to be transferred to the approved third party proteomic or genomic or service providers (such as QUB, Olink or GMI). In this event, all samples will be pseudo-anonymised. The codes relating to participant identification will not be disclosed to this study partner and remain in the custody of the chief investigator. Third party providers have already been approved as service providers with satisfactory evidence of encryption in storage and transfer of data and appropriate Material Transfer, Confidentiality and Intellectual Property agreements signed off by both the third party and the University. Unused DNA or plasma which is analysed by a third party such as GMI or Olink® proteomics will be destroyed. The research organisations that will use samples or data derived from analysis of samples could be not-for-profit e.g. universities or for-profit commercial companies (such GMI). GMI, however, have agreed to undertake*:* SNP genotyping, WGS and RNAseq generation analysis at no cost. Participant's name, address and personal details will not be made available to any organisation beyond the study team

Inclusion criteria

All cohorts:

- Gender: Male and female
- Age: >18
- BMI: Any
- Ethnic origin: Any

Hospitalised, non-hospitalised and nursing home patients:

- Symptomatic/non-Symptomatic patients testing positive or suspected of being positive for COVID-19/SARS-CoV-2 by virus PCR assay
- Nursing homes caring for elderly residents
- COVID-19 wards of participating hospitals

***Hospital and nursing home staff*:**

- Any member of hospital or nursing home staff are eligible to participate whether they have ever tested positive or not for COVID-19/SARS-CoV-2 by virus PCR assay even if non-symptomatic.

Exclusion Criteria

All cohorts:

- Under 18 years of age
- Those with intellectual disabilities or mental health illness

Blood sample collection

All venous blood samples will be taken by a trained phlebotomist or member of the research clinical team.

Data management

Enduring consent

The potential use of a participant's blood samples and data for future research studies is explained in the Participant Information Sheet and at the time of formal written consent. Participants will be advised they may choose (but are not obliged) to allow the use of their tissue in further studies (including or excluding genetic studies).

Participant confidentiality

Participants will be given a unique study code for identification purposes so that anonymity is protected.

Data records

The researchers will be responsible for running the study under the supervision of the Chief Investigator at Ulster University Participants will be assigned a unique study identification code. All electronic participant data will be pseudo-anonymised and be kept on encrypted and password protected university computers. Demographic and clinical or phenotypic data will be accessed from associated COVID-19 mobile device apps used at the time of sampling. One set of identifiable electronic data will be saved on the University password protected

server. Other data will be stored in locked cabinets under the custodial care of the chief investigator in accordance with the NHS Guidelines (HSC99/053). All genome sequence data that is stored on third party genomic service provider computers will be encrypted to ensure confidentiality. The third party will never have access to obvious personal identifiers, these are held solely by Ulster University for the purposes of follow up. Therefore, all biological and clinical data provided to the third-party provider is pseudo-anonymised and linked only by study identification code. Data encryption offers data security within and between Ulster and the third party. Encryption protects the personal identifiers which reside only in Ulster University and are never divulged outside. The encryption also protects data moving between organizations: 1) anonymised clinical data provided to the third party from Ulster University, and 2) anonymised biological data provided to Ulster University from the third party. Hard copies of the data will be kept under locked conditions, designed for the purpose. The Chief Investigator will act as custodian for all study data. Participants will not be identifiable in any data published in peer reviewed journals or in publicly accessible databases from this research. The dissemination of data will be carried out in agreement with the intellectual property arrangements that are in place for this research.

SARS-CoV-2 IgG or any other biological data will not be relayed back to participants. The data is for research use only and could not be supplied in this study as part of a clinical decision or back to work advisory.

All data would be deposited in archives such as The **G**lobal **I**nitiative on **S**haring All **I**nfluenza **D**ata **GISAID https://www.gisaid.org** (available for free both for deposition of data and download). Requests for access to data are protected through a data access committee).

Additional information about how we use data (in line with the General Data Protection Regulations (GDPR)

The General Data Protection Regulation (May, 2018) is an EU authority for the use of personal data, which has implications for health research. This new legislation requires that the researchers clearly describe the legal basis for, and how, they are processing study participants' data.

Ulster University is the sponsor for this study based at the Northern Ireland Centre for Stratified Medicine, Northern Ireland. Ulster University will be using participants' information in order to undertake this study and will act as the data controller for this study. This means that Ulster University is responsible for looking after the participant information and using it properly. Ulster University will keep identifiable information about participants for 10 years after the study has finished, unless participants have consented to being contacted about future studies.

Participants' rights to access, change or move your information are limited, as researchers at Ulster University need to manage participants' information in specific ways in order for the research to be reliable and accurate. If participants withdraw from the study, Ulster University will keep the information already obtained. To safeguard participants' rights,

Ulster University will use the minimum personally-identifiable information possible. (https://www.ulster.ac.uk/about/governance/compliance/gdpr).

Ulster University researchers involved in this study will use participants' names, clinical care number and contact details (telephone number and address/email address) to contact them about the research study, and make sure that relevant information about the study is recorded for their care, and to oversee the quality of the study. Individuals from Ulster University Governance and regulatory organisations may look at participants' medical and research records to check the accuracy of the research study. Ulster University researchers will pass these details to Ulster University Governance staff, along with the information collected from the participants and their medical records. The only people in Ulster University who will have access to information that identifies the participants will be people who need to contact them about appointments; invite them to be involved in other research projects, if the participant has given their permission, or; auditors of the data collection process. The people (outside the immediate research team) who will analyse the information will not be able to identify the participants and will not be able to find out their names, clinical care number or contact details. Ulster University will keep identifiable information about the participants from this study for 10 years after the study has finished, unless participants have consented to being contacted about future studies.

Ulster University researchers will collect information about the participants for this research study from their medical records in paper and electronic format. This information will include the participant's name, clinical care number, contact details and health information, which is regarded as a special category of information. Ulster University will use this information to generate a database to help the researchers answer the research questions they have identified in the Participant Information Sheets.
When the participants agree to take part in a research study, the information about their health and care may be provided to researchers running other research studies in this organisation and in other organisations. These organisations may be universities, NHS organisations or companies involved in health and care research in this country or abroad. Participants' information will only be used by organisations and researchers to conduct research in accordance with the UK Policy Framework for Health and Social Care Research. This information will not identify the participants and will not be combined with other information in a way that could identify the participants. The information will only be used for the purpose of health and care research, and cannot be used to contact the participants or to affect their care. It will not be used to make decisions about future services available to the participants, such as insurance.

Statistical Analysis

A statistical sample size calculation was based on the core objective of measuring SARS-CoV-2 infection prevalence in specified populations: at community level (nursing home), in hospitalised SARS-CoV-2 positive cases and in frontline healthcare staff. If the true

prevalence of COVID-19 in frontline staff is 0.20 (20%), with a sample size of n = 246 we are 95% confident that our estimate will not be more than 5% (desired precision) below or above the true prevalence (20%). If we add 20% contingency in sample size for loss of follow up or drop out, we estimate that n = 295 frontline staff from each setting would be required. In another scenario, If the true prevalence of COVID-19 in hospitalized patients and nursing home residents is 0.60 (60%), with a sample size of n = 369 we are 95% confident that our estimate will not be more than 5% below or above the true prevalence (60%). Again, allowing for 20% contingency, we estimate that n = 443 patients from either setting are required.

For the remaining core objectives, a predictive model would be built and tested for the severity of COVID-19 in infected individuals. Multi-predictor variable regression modelling would be the appropriate technique to achieve these. Disease severity will be the primary dependent variable of interest, with age, comorbidity, gender, immune response measure, virial strain, and host genetics as candidate predictor variables, although some of them may turn out to be non-significant predictors. Therefore, we anticipate 5-predictor models to be constructed from the dataset. Based on the Guideline of TRIPOD on the EPV (events per predictor variable), we used the following to determine the appropriate samples size i.e. 4 levels of severity measure X 20 events per predictor X 5 predictors = 400. Allowing for some loss of sample due to follow up we estimate that recruiting a sample of n=500 total infected individuals from both settings is sufficient for our study. The 20 EPV is considered to be a small to medium dataset and it is well above the minimum 10 EPV recommended for building predictive models.

Standard approaches will be used to detect patterns in missing data. Data will be modelled using bespoke packages (such as DADA2 for 16S rRNA, or RnBeads for epigenetics) in R, or using approaches in Galaxy, tools from Bioconductor or similar. Data exploration approaches such as PCA and hierarchical clustering will be used to look for trends and possible confounders in the data. Post-treatment outcome data will be summarised both in graphical and tabular form for each metric. Patients will be assigned non-response or good response status based on primary outcome measures. The data will be analysed to examine any relationships between host/viral genomic data, OLINK, CRP, flow cytometry immune cell data and outcome measures. Other established statistical hypothesis testing methods, including t-tests, ANOVA, linear models, or non-parametric tests depending on the distributions of the measures observed will be used to analyse secondary outcomes

References

Merad M, Martin JC. Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. Nat Rev Immunol. 2020 May 6:1–8. doi: 10.1038/s41577-020-0331-4. Epub ahead of print. PMID: 32376901; PMCID: PMC7201395.

Moons KG, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1-W73.

Ong EZ, Chan YFZ, Leong WY, Lee NMY, Kalimuddin S, Haja Mohideen SM, Chan KS, Tan AT, Bertoletti A, Ooi EE, Low JGH. A Dynamic Immune Response Shapes COVID-19 Progression. Cell Host Microbe. 2020 Apr 30:S1931-3128(20)30185-2. doi: 10.1016/j.chom.2020.03.021. Epub ahead of print. PMID: 32359396; PMCID: PMC7192089.

Penny Ward. COVID-19/SARS-CoV-2 Pandemic 2020.https://www.fpm.org.uk/blog/covid-19-sars-cov-2-pandemic/
https://www.hdruk.ac.uk/projects/national-phenomics-resource/

Sun X, Wang T, Cai D, Hu Z, Chen J, Liao H, Zhi L, Wei H, Zhang Z, Qiu Y, Wang J, Wang A. Cytokine storm intervention in the early stages of COVID-19 pneumonia. Cytokine Growth Factor Rev. 2020 Apr 25:S1359-6101(20)30048-4. doi: 10.1016/j.cytogfr.2020.04.002. Epub ahead of print. PMID: 32360420; PMCID: PMC7182527.

Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. Nat Rev Immunol. 2020 Apr 28:1–12. doi: 10.1038/s41577-020-0311-8. Epub ahead of print. PMID: 32346093; PMCID: PMC7187672.

Thevarajan I, Nguyen THO, Koutsakos M, Druce J, Caly L, van de Sandt CE, Jia X, Nicholson S, Catton M, Cowie B, Tong SYC, Lewin SR, Kedzierska K. Breadth of concomitant immune responses prior to patient recovery: a case report of non- severe COVID-19. Nat Med. 2020 Apr;26(4):453-455. doi: 10.1038/s41591-020-0819-2. PMID: 32284614; PMCID: PMC7095036.

Vardhana SA, Wolchok JD. The many faces of the anti-COVID immune response. J Exp Med. 2020 Jun 1;217(6):e20200678. doi: 10.1084/jem.20200678. PMID: 32353870; PMCID: PMC7191310.