# Protocol

Protocol for: Hall V, Foulkes S, Insalata F, et al. Protection against SARS-CoV-2 after Covid-19 vaccination and previous infection. N Engl J Med. DOI: 10.1056/NEJMoa2118691

This trial protocol has been provided by the authors to give readers additional information about the work.

This supplement contains the following items:

# SIREN: The impact of detectable anti SARS-COV2 antibody on the incidence of COVID-19 in healthcare workers

Synopsis

| | |
|---|---|
| Study title | Impact of detectable anti-SARS-COV2 on the subsequent incidence of COVID-19 in healthcare workers |
| Short title | SIREN (**S**arscov2 **I**mmunity & **RE**infection Evaluatio**N**) |
| Chief Investigator | Susan Hopkins |
| Sponsor | Public Health England |
| Funder | Department of Health and Social Care |
| Study coordination | Public Health England |
| Statistics and design | Andre Charlett |
| Study management group | Study design and management: Susan Hopkins, Meera Chand, Colin Brown<br>Statistics: Andre Charlett<br>Epidemiology: Mary Ramsay<br>Virology and serology: Maria Zambon, Tim Brooks<br>Data management: Annie Marie O'Connell<br>Genomics: Richard Myers<br>Information governance: Samantha Organ |
| Study Design | Prospective cohort study to determine incidence of new SARS-COV2 infection in seronegative and seropositive healthcare workers |
| Study participants | NHS staff working in clinical settings in England |
| Sample size | 10,000 healthcare workers |
| Planned study period | 24 months |
| Planned recruitment period | 3 months with follow-up for up to 18 months; interval analysis of incidence and prevalence monthly and statistical analysis on primary outcome 3 monthly. |

## Summary

COVID-19 is causing a global pandemic. This study aims to find out whether healthcare workers who have evidence of prior COVID-19, detected by antibody assays (positive antibody tests), compared to those who do not have evidence of infection (negative antibody tests) are protected from future episodes of infection.

In this study, we will recruit healthcare workers to be followed for at least a year and study their immune response to the virus causing COVID-19, called SARS CoV2. We will do this by collecting data on their history of COVID-19 infection and any new symptoms. All NHS staff who deliver care to patients are being asked to have a nose and throat swab every other week in order to detect mild cases or cases who do not have symptoms. This is the main test that is currently used to detect and diagnose infection. It looks directly for the virus in the nose and throat. Once the infection is cleared, we cannot detect virus in samples. Therefore, we will also ask these individuals to have blood samples taken every other week to determine whether they have antibodies to the infection. These blood samples allow the previous infection to be detected as the response to infection in the body is to produce small particles in the blood called "antibodies". It takes up to 4 weeks to make enough antibodies to fight the infection. But once someone recovers, antibodies stay in the blood at low levels– this is may help prevent us from getting infected with the same infection again. However, for SARS CoV2 infection we do not know yet if the detection of antibodies protects people from future infection. Through this study we will provide this very important information which will help to understand the future impact of COVID-19 on the population.

## 1.0 Background

On 31st December 2019, the first cases of infection with a novel coronavirus, subsequently designated SARS-CoV-2, emerged in Wuhan, China.[1] A global pandemic was declared by the World Health Organisation (WHO) on 12th March 2020.[2] By 9th May, there were more than 3.6 million confirmed infections and 240,000 reported deaths globally.[3]

In order to determine the population impact of potential future waves of the pandemic of COVID-19, there is an urgent need to understand whether prior infection and any host immune response provides protection from future reinfection with SARS-CoV2.

Currently, a confirmed diagnosis of infection of SARS-CoV2, relies on laboratory diagnosis of infection using real-time reverse transcriptase polymerase chain reaction (RT-PCR). Diagnostic RT-PCR typically targets the viral ribonucleic acid (RNA)-dependent RNA polymerase (RdRp) or nucleocapsid (N) genes using swabs collected from the upper respiratory tract (nose and throat).[4, 5] However as RT-PCR only detects acute infection we are unaware of the true population prevalence of infection. While there is increasing evidence that a proportion of the population can be infected and asymptomatic, the true proportion of asymptomatic infection in the population is unknown.

In England, there is a new national NHS policy to screen healthcare workers with RT-PCR at least every other week to determine the incidence of new infection and reduce the risk of healthcare worker (HCW) transmission to other HCW and patients. However, RT-PCR from upper respiratory tract swabs may be falsely negative, due to quality or timing of collection; or due to viral titres in upper respiratory tract secretions peaking in the first week of symptoms,[6] but declining below the limit of detection in patients who present with symptoms beyond this time frame.[7] In individuals who have been infected and recovered, RT-PCR provides no information about prior exposure or immunity. In contrast, assays that reliably detect antibody responses specific to SARS-CoV-2 could contribute both to diagnosis of acute infection (via rises in IgM and IgG levels) and to identify those who have been exposed and recovered with or without symptoms (via persisting IgG).[8,9] A further research gap is whether the detection of an antibody response to SARS-CoV2 will provide immunity to re-infection. In a study of medical students, reinfection with HCoV-229E was detected in 937 student years of follow-up; neutralizing antibody to 229E was commonly but it did not appear to influence the occurrence of, or likelihood of illness with, reinfection as judged by complement fixation seroconversion.[10] While IgG and IgM to SARS-CoV2 can be effectively used to determine past exposure. It may be that these antibodies do not protect against future infection. The presence of **neutralising antibodies** may be required as these not only confirm prior infection but also prevent future infection. There are no commercial assays available to detect neutralising antibodies at present but PHE and other university partners are developing in house assays to detect these specific antibodies. We will examine correlates of different types of antibody response and the protective effect against re-infection.

Therefore, this study will answer the key question on whether prior SARS-CoV-2 infection confers future immunity to SARS-CoV2 re-infection in health care workers who have a higher prevalence of infection than the general population; in one study 20% of asymptomatic London HCW tested for SARS-CoV2 were positive on RT-PCR tests over a 5-week period from 23rd March to May 3rd, with incidence peaking at 7.1% and declining 6 fold after 5 weeks.[11] It will also determine whether reinfection is due to a lack of neutralising antibody or waning immunity over time in individuals who are anti-SARS-CoV-2 positive, whether such individuals can be re-infected symptomatically or asymptomatically, about the incidence of new infection in individuals without prior exposure and the severity of COVID-19 in initial and subsequent episodes of infection if re-infection occurs. We will also use genomics to determine the viral diversity in healthcare workers at each sampling point and particularly if individuals are re-infected.

## 2.0 Aim and Objectives

Aim: The overall aim of this study is to determine if prior SARS-CoV-2 infection in health care workers confers future immunity to re-infection.

Objectives:

Primary: To determine whether the presence of antibody to SARS-COV2 (anti-SARS-COV2) is associated with a reduction in the subsequent risk of re-infection over the next year

Secondary:

1. To estimate the prevalence of SARS-CoV-2 infection in healthcare workers by region, using baseline serological testing at study entry of healthcare workers and symptom history from February 1st 2020 to date of study entry
2. To estimate the subsequent incidence of symptomatic and asymptomatic SARS-CoV-2 infection and determine how this varies over time, using fortnightly RT-PCR testing (combined with any intercurrent testing on development of symptoms)
3. To estimate cumulative incidence of new infections in healthcare workers stratified by age, sex, ethnicity and co-morbidities.
4. To measure the ability to culture viable virus from cases of re-infection diagnosed by RT-PCR and whether those that are persistently positive on RT-PCR are continuing to shed viable virus.
5. To use genomic comparison to determine whether healthcare workers who become PCR-positive for a second time within a defined time frame are experiencing persistent infection or re-infection.
6. To determine whether serological response wanes over time
7. To determine whether there is a relationship between serological response (using enzyme immunoassay detection of IgG) and the presence of neutralising (protective) antibodies
8. To identify serological, demographic or clinical factors that correlate with the presence of neutralising antibodies
9. To investigate the phylogenetic relatedness of SARS-COV2 viruses causing healthcare worker infections
10. To investigate the relationship between illness severity, demographics and neutralising antibody production

# 3.0 Methods

## 3.1 Study design

The overall study design is a prospective longitudinal cohort study of NHS staff at representative acute Trusts across England

## 3.2 Study sites

Each acute Trust will be asked to recruit invite all staff to participate, with allocated quotas within the cohort at each site to ensure appropriate pro rata representation across doctors, nurses, professions allied to medicine, healthcare assistants, administrative staff in clinical settings, and porters i.e. professions which have a patient facing role. A total of 40 NHS hospitals will be asked to participate; five hospitals in each of the six NHS regions outside London and oversampling of 10 Trusts in London given the higher incidence of disease, higher density of the population, and higher likelihood of seeing re-introduction of cases.

## 3.3 inclusion/exclusion criteria

### Inclusion criteria

- Health care worker with a role providing direct patient care
- Can provide written consent
- Is likely to be remain engaged with follow up, even if they move to another hospital or Trust site

### Exclusion criteria

- Written consent not provided
- Temporary staff member

## 3.4 Sample size

A simulation approach has been used to estimate the power to detect relative differences between the study cohorts.

It is assumed that a total of 40 acute Trusts will be selected for this study. It is anticipated that in total 10,000 healthcare workers will be recruited; and with an estimated 25% seropositivity of healthcare workers (based on 20% of staff who were asymptomatic and tested positive in one London hospital between 23rd March and 2nd May 2020[11]), this would provide a cohort of 2,500 seropositive HCWs and 7,500 seronegative HCWs to be recruited to this study.

It has been assumed that on average 250 healthcare workers would be recruited from each selected Trust, with a standard deviation of 50. The proportion of seropositive recruits at each Trust has been obtained from a Gaussian distribution with a mean of 0.25 and standard deviation of 0.05 to reflect expected inter-Trust variation. The cumulative incidence in each Trust in the seronegative cohort has been simulated using Gaussian distributions with means of 0.05, 0.1, 0.2 and 0.3 each with a coefficient of variation of 0.2. This range represents that which is feasible to observe over a 12-month period, given the behavioural and social interventions still being employed during the study to control transmission.

A study duration of 52 weeks has been assumed with the inter-test period of 2 weeks. A total attrition of 35% of HCWs is assumed, unaffected by sero-status and occurring at a constant rate over the 52 weeks of the study. Attrition has been assumed to be independent of the infection process.

A range of cumulative incidence in the seronegative cohort has been used to reflect the immune effectiveness; with units in the simulations being allocated to be infected or not, using a draw from a Bernoulli distribution with p equal to the Trust and cohort specific simulated cumulative infection rate. A simplifying assumption of a constant infection rate over the study period has been used.

For each scenario as set of 200 simulations were performed. For each simulation, the total number if infections and person weeks of follow up was calculated for each cohort in each Trust. This data was analysed using a mixed effects Poisson model, using the natural logarithm of the person weeks as an offset. Power was estimated as the proportion of simulations for which the Wald statistic p value for the estimated incidence rate ratio of the seropositive to seronegative cohorts was less than 0.05. These are

presented in Table 1, indicating that there is sufficient power for all but the smallest immune efficacy of 0.1 i.e. a 10% reduction in incidence in the seropositive cohort. Such a small reduction is indicative of an irrelevant level of protection to provide a means of controlling the pandemic via natural herd immunity.

**Table 1: Power estimates obtained via simulation for a range of immune effectiveness and cumulative incidence**

| Cumulative incidence in the seronegative at baseline cohort (per 100 HCW) in 12 months | Immune Effectiveness 10% | Immune Effectiveness 20% | Immune Effectiveness 30% | Immune Effectiveness 40% | Immune Effectiveness 50% |
|---|---|---|---|---|---|
| 0.05 | 0.15 | 0.44 | 0.79 | 0.98 | 1.00 |
| 0.1 | 0.20 | 0.77 | 0.99 | 1.00 | 1.00 |
| 0.2 | 0.53 | 0.99 | 1.00 | 1.00 | 1.00 |
| 0.3 | 0.67 | 1.00 | 1.00 | 1.00 | 1.00 |

To investigate the ability of a study of this size to detect differential immune efficacy in subgroups of healthcare workers, for example those aged over 50s or from a BAME group a further set of simulations were undertaken. It has been assumed that the subgroup of interest is on average 35% of the total cohort size. The cumulative incidence used in the previous simulations has again been used with just those in the subgroup that are seropositive having an absolute reduction in immune efficacy of 5%, for example, if the cumulative incidence in the seronegative cohort is 0.3 (30%) and the immune efficacy in the non-sub group is 0.5 with a cumulative incidence of 0.15 (15%) in this group, the cumulative incidence in the seropositive sub group is 0.2 (0.15+0.05 20%). The interaction between cohort type and subgroup has additionally been added to the model and the estimated power obtained using the proportion of the 200 simulations for which this interaction had a Wald test p value of less than 0.05. Only immune efficacies of 0.3 or greater have been used, and these are presented in Table 2.

**Table 2: Power estimates obtained via simulation for a range of immune effectiveness and cumulative incidence for detecting a sero-status by subgroup interaction**

| Cumulative incidence in the seronegative at baseline cohort (per 100 HCW) in 12 months | Immune effectiveness 30% | Immune effectiveness 40% | Immune effectiveness 50% |
|---|---|---|---|
| 0.1 | 0.93 | 0.91 | 0.98 |
| 0.2 | 0.70 | 0.82 | 0.84 |
| 0.3 | 0.67 | 0.68 | 0.69 |

## 3.5 Site set up

A site initiation visit (SIV) will take place, in person or by telephone, for each study site. During the SIV, a member of the central study team will describe the study methods and discuss with the clinical and research team possible ways to organise recruitment and data collection. An electronic recruitment database will be provided to be completed by the site including all data PHE requires for data linkage and for monitoring the cohort composition. This will include the NHS number, name, date of birth, gender, date of enrolment, occupation, and study number. Study sites will be responsible for securely transferring an updated extract of this data base to PHE each week.

## 3.6 Recruitment

Each Trust will be provided with a timeline by which enrolment must be complete. Participants are recruited by an all-staff communication requesting volunteers and offering open sessions with members of the local study team.

Screening, eligibility assessment and formation of the cohort at each hospital: For each volunteer, a member of the local study team will review eligibility assessment. The proportions of participants from different occupational groups will be monitored. The cohort will be formed by accepting consecutive eligible volunteers until the quota for each staff group is filled. The site team may choose to send out further communications to support recruitment to under-represented groups. If a particular group's quota cannot be filled, those places may be used for staff from other groups. The study is observational only and there is no randomisation or allocation to any subgroups.

Staff such as junior doctors may have planned moves to other Trusts within the study period. For these staff, they will be asked to continue to complete the questionnaire and either attend the enrolment hospital for regular swabs and bloods or transfer this function to their new hospital if possible.

All formal consent procedures will be undertaken by the team in each trust and participant will give full informed written consent.

## 3.7 Data collection and Measurements

### 3.7.1 Baseline data

**Self-completed baseline questionnaire**: The participant will self-complete an online survey at enrolment. Baseline data will include demographic/social factors (gender; date of birth; ethnicity; Smoking history); comorbidities including whether immunocompromised; job role including exposure to aerosol-generating procedures; whether had laboratory confirmed COVID-19 since February 1st, 2020, or symptoms compatible with COVID-19. (Appendix 1)

**Initial serological assessment**: The participant will have a blood taken at baseline which will be separated and aliquoted locally at each Trust. One aliquot will be tested locally for anti-SARS-CoV-2 using the commercial assay in use routinely in the trust. This will be used to assign the participant to the seronegative or seropositive cohort at baseline. The second aliquot will be stored at -70$^{O}$C and batch shipped to PHE for later analysis using alternative assays to establish a correlate of protection and compare the serological responses across different assays.

### 3.7.2 Ascertainment of Outcome Measures

**Frequency of follow up visits:** If possible the site team at each Trust will coordinate each study follow up visit to coincide with the routine PCR screening process being undertaken for all healthcare workers in the Trust. At the initiation of the study it is envisaged this will occur every two weeks; frequency may be altered to be more or less frequent (1 - 4 week intervals) depending on national and regional epidemiology.

**RT-PCR screening**: This will be undertaken according to the local Trust procedure for screening healthcare workers. It may be an administered nose and throat swab, or a self-swabbing kit depending on local trust protocol. The swab will be tested for SAR-CoV-2 RNA in an accredited laboratory, which may be the clinical laboratory or through the national testing programme, using a reverse transcriptase polymerase chain reaction (RT-PCR). As this is a clinical sample that requires reporting of the result to each individual HCW, full standard identifiers will be used.

**Anti-SAR-CoV-2 testing**: Each participant will have a venous blood test on the same day as the RT-PCR screening (within 48 hours is acceptable). The site team will determine the optimal procedure within each Trust, which may be visiting a study nurse, attending the phlebotomy service with a pre-designed form, or similar. A 10 ml venous blood sample will be taken. This will be separated in the local clinical or clinical research laboratory, the pellet discarded, and as many 1 ml aliquots of serum as possible will be formed. One aliquot will be tested immediately using the enzyme-linked immunosorbent assay (ELISA) rolled out as part of the national serological programme for healthcare workers. As this is a clinical sample that requires reporting, full standard identifiers will be used. A further aliquot will be sent to PHE for multi-assay testing in non-real time. **Reporting of results:** The local PCR and serological results will be reported according to the Trust's standard occupational health procedures.

**Action on positive results**: If the PCR screen is positive, the participant will follow the national guidance for self-isolation and the Trust guidance for return to work. If the serology assay is positive, the participant will receive advice indicating that this does not mean they are immune to reinfection, and that they must continue to adhere to infection control measures at home and work as usual.

**Follow up questionnaire**: The participant is asked to log in to the online survey system and fill in the monitoring questionnaire at the same frequency and timing of the RT-PCR and Anti-SAR-CoV-2 testing. The follow up questionnaire will collect information on symptoms consistent with COVID-19 since last study visit, date of onset, duration, laboratory results, contract with definite case of Covid-19 (Appendix 2). This will be identified by their unique study identifier only.

**3.8 Participant withdrawal**

Participants may withdraw at any time and this is explained in the participant information leaflet. On a request to withdraw they will be given the following options:

- Withdraw from any future study visits, but allow their existing materials and data to remain in the study and be tested, and allow the study team to continue to access their results from their routine PCR testing undertaken by the Trust.
- Withdraw from any future study visits, and withdraw the use of their existing materials, but allow the data generated up to the date of withdrawal to remain in the study. Their samples will be destroyed but their record will remain in the study database.
- Withdraw from any future study visits, and withdraw the use of their existing materials, and have their data removed from the study. Their samples will be destroyed and their record removed from the study database.

### 3.9 Storage of materials and additional Laboratory testing

All residual swab material, nucleic acid extract, and serum will be stored at the hospital site and shipped to PHE in monthly batches. They will be used for confirmatory testing or for serological or viral genomic characterisation. No human DNA genomic investigations will be undertaken. Additional serological testing: Seropositive participants in whom reinfection is identified, plus a cohort of matched non-infected seropositive controls, will have their sera further characterised for the presence of neutralising antibody, to provide hypothesis generating data on mechanisms of protective immunity.

**Genomic analysis**: All positive samples from participants will be sequenced as part of the routine sequencing of NHS residual samples in COG-UK Consortium laboratories. For participants who have more than one PCR test positive, genomes will be compared where possible to provide evidence to support reinfection or persistent infection as a mechanism. Phylogenetic analysis of SARS-COV2 from healthcare workers, using the study samples and the wider collection of genomes available through the COG-UK Consortium, will also be undertaken as an exploratory analysis into the diversity and spread of SARS-COV2 in healthcare workers.

## 4 Data management

All study documentation will be stored at each site, either in hard copy in a secure environment and/or in electronic copy in an access-limited location on a Trust server, as decided suitable by the Trust. Information received by PHE as the central study site will be stored securely in access-limited locations on PHE servers.

Study data is as follows:

- Recruitment log, containing personal and demographic information, stored by the Trust
- NHS and Lighthouse laboratory records, identified by personal information – this data (including name, date of birth and postcode) is automatically transmitted to PHE as part of the official notification of infection (under Health Protection Regulations) – it includes both results of RT-PCR tests and serology results performed by the laboratories. The study will access the routine data generated from the national screening programme in the participant Trusts with no additional laboratory requirements. The data collected by PHE will include the sample type, assay used and cycle threshold value for each sample tested.
- Questionnaire data entered by the participants (pseudonymised, with the link to personal information held in the recruitment log). This will be managed using SnapSurvey, a data management system held on a PHE server with end to end encryption.
- Viral genomic data generated by the local or consortium laboratories, identified by the Consortium identifier (pseudonymised with the link to personal information held by PHE under existing surveillance protocol)

All source data will be securely transferred to PHE with the identifiers described above. The transfer will be by secure email (recruitment log), through PHE's established Second Generation Surveillance System (SGSS; laboratory data), or through end to end encrypted bespoke arrangements (questionnaire and genomic data). The person information is required for PHE to undertake secure data linkage across the sets as well as drawing in data from the Lighthouse laboratories. The database containing personal information will be access-limited to essential PHE staff working in the study team. Once linked to a participant, the data will be pseudonymised using the participant's study identifier and stored on a mirror database containing pseudonymised data only. All analysis will be performed by study personnel on the pseudonymised linked datasets.

Access to Data: Direct access will be granted to authorised representatives from the Sponsor for monitoring and/or audit of the study to ensure compliance with regulations. All researchers involved in data linkage have been trained in handling data according to Caldicott guidelines and Section 60 of the Health and Social Care Act. All researchers are aware of the Data Protection Act 1998 and the need to maintain absolute confidentiality.

Data Recording and Record Keeping: The data will be securely held at the National Infection Service, PHE. Data collection, storage and use will be consistent with the procedures described in the NHS Information Governance Toolkit. All databases will be encrypted and appropriately access-restricted.

Data storage: Electronic data will be stored on PHE secure servers and will remain active for the duration of the study. Participants' identifiable data will then be removed and the data will be archived within the platform to be retained for a period of 5 years.

## 5. Analysis plan

All enrolled participants will be included in analyses, which will account for clustering by Trust. Analyses will be conducted after each 4-week period to inform the UK's response to the COVID-19 pandemic. Results will be available to all organisations involved in the research. The study will end by default after 18 months, but by consensus of the study management group and funder may be terminated sooner if findings are sufficient.

There are no formal stopping rules for futility, utility or lack of power. The final decision to terminate the study will be made by the Public Health England and Department for Health and Social Care.

Estimates of both cumulative incidence and incidence density in the seropositive and seronegative cohorts will be obtained using mixed effects models assuming counts of PCR positive have a negative binomial distribution, a log link function, and the natural logarithm of the total number of subjects or the total follow-up time use as an offset, respectively. Inclusion of a binary predictor indicating the sero-status of the cohort into this model will provide estimates of the incidence rate ratio. Trust will be incorporated as a random intercept to account for unmeasured, shared, Trust level factors. To account for a non-constant force of infection, calendar month will be incorporated as an additional random effect. An assessment of the role of factors such as age, gender, ethnicity in immunity will be explored by inclusion of interactions within the model between each and serological status.

While the above analytical approaches provide a “classical” person-years approach to prospective cohort analysis and provide familiar measures of association, it may be inadequate to assessment of immunity provided by seroconversion. As it is expected that seropositivity is likely to confer a degree of short to median term protection for a SARS-CoV-2 infection, multi-state and parametric cure rate models incorporating frailty will be employed. These “survival” type of models provide a more detailed assessment of factors associated with both short term and longer-term protection from infection, and how immunity may wane over time. Both mixture, explicitly assuming an immune and non-immune group and non-mixture “cure rate” models will be assessed using information criterion to choose which provides a better fit to the observed data. Bayesian approaches to cure rate models with frailty as describe by deSouza[17] will be employed.

Multi state models explicitly allowing those within an “immune” state to flow into a “susceptible” state as antibodies wane will also be employed. This framework can allow subjects to move from seronegative (susceptible) to seropositive (immune) when infected during the study period. An additional absorbing state will be used for those infected that died. It is also possible to introduce “misclassification” of state into the multi state model, providing an estimate of sensitivity to account for imperfect serological tests. Approach like those proposed by Jackson[18] will be employed.

Procedure for Accounting for Missing, Unused, and Spurious Data: Analyses will be restricted to complete cases. The RT-PCR test for virus is being used as a diagnostic test and hence has extremely high performance. Sufficient sera will be obtained to re-run the immunological assays in case of initial assay failure. For similar reasons we do not anticipate that spurious data will be obtained.

Procedures for Reporting any Deviation(s) from the Original Statistical Plan: Deviations from the original statistical plan or the statistical analysis plan will be described and justified in the analysis reports.

## 6. Confidentiality and information governance

The data management plan (Section 4) demonstrates that PHE will need to be able to extract data from multiple Trusts, including staff moving between Trusts, national testing laboratories undertaking clinical diagnostic tests when staff are unwell, and potentially national sequencing consortium laboratories. Data linkage based on a single study identifier is neither practical nor robust in such circumstances. The suggested process for handling identifying data in a secure and appropriate way is therefore as follows:

- On enrolment, participants will be given a unique study identifier (three letters identifying the Trust, plus five digits identifying the participant)
- The Trust will compile a participant database containing the name, date of birth, NHS number, and study identifier of each enrolled participant
- The Trust will retain this for local study management and will transfer a copy of the database securely to PHE using appropriate encryption and will be stored in a limited-access environment on a PHE server. The database will be updated and mirrored securely to PHE every month.
- PHE will use the participant's personal information to retrieve their test results from the national surveillance systems as described in the data management plan. They may also use personal information to request the residual samples for additional processing in the sub studies.
- Positive swabs from the study will be routinely sequenced under the arrangement between the NHS and the COG-UK Consortium. In order to retrieve the genomic data, PHE may use the participant's identifying information to search the COG-UK data hub, which is **within** PHE, and retrieve the COG-UK study ID. The COG-UK study ID allows PHE to retrieve the viral genomes from that participant's sample from COG-UK (which maintains a pseudonymised genome collection on a non-PHE server). SIREN study team will make a request to the COG-UK team and will not be able to access any other fields in the COG-UK data hub or access the data hub directly.
- Once PHE has retrieved the data from the appropriate systems, it will be uploaded to a study database identified **only** by the study identifiers, and with all participant identifying information removed.
- This database will be submitted monthly for analysis.
- If a participant moves to another Trust, PHE will be notified through the monthly update of the study log.

## 7. Quality assurance and research governance

The study may be monitored or audited in accordance with the study protocol and standard operating procedures, GCP and relevant regulations.

**Risk assessment:** No formal risk assessment is required. The study involves recruiting individuals without symptoms who will be asked to give full informed consent to have swabs taken or self-swab their throat and nose and provide a blood sample taken by a study research nurse or the hospital phlebotomy department. Further participation in follow up visits to collect the same samples is based on consent of the participant. The main burden of participating in the study is the time taken for the study visits (which will be as far as possible at the same time as routine trust mandated screening) or the potential for minimal bruising from blood sampling though this is unlikely with experienced staff taking the samples. There is minimal risk of harm to any patient from participating since it does not include any therapeutic intervention. The surveys do not ask personal or intrusive information. The diagnostic test for the presence of virus from the nose and throat swab will be conducted by an accredited laboratory and will be returned to the participant's research nurse and HCW through standard hospital practice.

**Study monitoring:** No GCP monitoring will be undertaken. As described there are minimal risks posed to patients by this observational and non-interventional study. The only study procedures are the completion of the questionnaire and taking of samples.

**Safety reporting:** There are no interventions in this study, and the only procedures are a standard blood draw performed by study HCP and a participant self-swab using a methodology that is being used widely across the country. Therefore, there is minimal safety risk to participants and safety reporting is not applicable.

**Study committees:** Oversight will be provided by a study management group including the investigators named above, representatives of collaborating and participating organisations as appropriate, and chaired by the Chief Investigator.

**Protocol deviations**: A deviation is a departure from the approved study protocol or other study document or process, or from Good Clinical Practice or any applicable regulatory requirement. Any deviations from protocol will be documented in a protocol deviation form and filed in the study master file.

**Serious breaches**: A serious breach is a breach of the protocol or of the conditions or principles of Good Clinical Practice which is likely to affect to a significant degree –

(a) the safety or physical or mental integrity of the trial subjects; or

(b) the scientific value of the research.

In the event that a serious breach is suspected the Sponsor must be contacted within 1 working day. In collaboration with the Chief Investigator, the serious breach will be reviewed by the Sponsor and, if appropriate, the Sponsor will report it to the approving REC committee and the relevant NHS host organisation within seven calendar days.

# Appendix 1: Participant documents

PARTICIPANT INFORMATION LEAFLET

SIREN- SARS-COV2 immunity and reinfection evaluation

Impact of detectable anti SARS-COV2 antibody on the incidence of COVID-19 in healthcare workers

**INFORMATION LEAFLET FOR PARTICIPANTS**

We would like you to invite you to take part in this study to understand whether prior infection with SARS-CoV2 (the virus that causes COVID-19) protects against future infection with the same virus. *Before you decide, it is important that you understand why the research is being done and what it would involve for you. Please take time to read this information, and discuss it with others if you wish.* Please also ask the research nurse if there is anything that is not clear.

**Why are we doing this study?**

The coronavirus (COVID-19) pandemic is having a major impact across the UK. This study aims to find out if those who have had the infection in the last three months and if you have been infected already whether this protects you against future infection with this virus.

One way to find out whether a person has an infection is to directly look for the virus in their nose and throat. COVID-19 is caused by a virus, and the main test we are using to diagnose it at the moment is a test to find this virus. Once someone has recovered from the infection, the virus is no longer present in the nose or throat. But one way the body fights infections like COVID-19 is by producing small particles in the blood called “antibodies”. It takes 2-3 weeks for the body to make enough of these antibodies to fight the infection. When someone gets better, these antibodies still stay in their blood at low levels – this can help protect against future infections with the same virus. In this study we will measure active infection and also measure who has had previous infections using antibodies to their blood.

By doing both these tests regularly together over time we will be assessing whether prior infection (measured through an antibody test) protects against future infection. You will be given all your results as they are performed at your local laboratory and the results will be shared with PHE so that they can measure the impact over time. By taking regular samples, we can measure what proportion of frontline NHS staff have been exposed to SARS-CoV-2 and how quickly SARS-CoV-2 is spreading over the coming months. The donated samples will be treated as a gift meaning that we will not be able to return them to you.

**Why have I been asked to take part?**

Taking part is voluntary and you should not be placed under any pressure - it is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to

keep and asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason.

**What do you want me to do?**

If you decide to take part, you will be asked to complete a registration questionnaire and provide an updated symptom review, nose and throat swabs, blood sample every two weeks for up to 12 months.

If you have fever, cough or any other respiratory symptoms or you are a contact of a confirmed case of COVID-19 are currently being asked to self-isolate at home, then please access swabbing and complete as you would normally within your organisation. If that is the case, please reschedule your appointment for a later date when you are well and back at work

**What will happen to my sample?**

Your sample will be processed as normal to look for active virus from nose and throat swabs; if you are asymptomatic and test positive then you will be asked to stay off work for at least 7 days, or at least 7 days from the time you develop symptoms if those occurred just before or within the 7 days after you have had a positive test.. The blood samples will be processed in a laboratory to collect the serum and the part containing the majority of cells will be discarded. Your blood samples will be tested for antibodies against SARS-CoV-2. Testing may be done at a PHE reference laboratory, at your local hospital or both. Any remaining serum sample at the end of the survey will be anonymised and incorporated into the PHE Seroepidemiology Unit (SEU) collection, unless you ask us to destroy your sample as soon as it has been tested. Samples stored at the PHE SEU will be used to perform a range of different national serosurveys in the future. If you do not want us to transfer your sample to the PHE Seroepidemiology Unit collection, this will not prevent your taking part in the surveillance, this is marked on the consent from. Your personal data will be stored in accordance with the General Data Protection Regulations (GDPR) and the Data Protection Act 2018.

**What are the benefits to me?**

The study will not benefit you directly, but your participation will help provide important information about SARS-CoV-2 re-infection among clinical NHS healthcare staff and provide a stronger evidence base to inform national guidance and policy. You can also ask to be informed of your SARS-CoV-2 antibody results when they become available. At the end of the surveillance, the overall results will be published in national reports.

**What are the disadvantages?**

For some, blood sampling may cause momentary discomfort at the site of the blood draw, possible bruising, redness, and swelling around the site, bleeding at the site, feeling of lightheadedness when the blood is drawn, and rarely, an infection at the site of blood draw

**What if I change my mind?**

If you no longer want to be involved, you can withdraw from the survey at any time by contacting us (contact details below).

**Will I be given my results?**

Yes, you can be given your results according to your local Trust procedures. If you have SARS-COV2 detected on your nose/throat swab, you must follow the self-isolation guidance as you would usually. If you have SARS-COV2 antibody detected in your serological test, you should remember that this does not necessarily protect you against future infection, and you should not change your behaviour. You should take all usual precautions against COVID-19 at home and at work.

**What should I do now?**

If you would like to volunteer, please complete and sign the consent form. If you have any questions regarding the study, please contact us (contact details below)

**Who has reviewed the study?**

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, to protect participants' interests. This study has been reviewed and given favourable opinion by _____Research Ethics Committee.

**How have healthcare workers been involved in the study?**

This study has been reviewed by healthcare workers who have participated in PHE incidence (through swabbing) and serosurveillance studies.

**What should I do if I have any concerns?**

Public Health England as Sponsor, has appropriate insurance in place in the unlikely event that you suffer any harm as a direct consequence of your participation in this study.

If you wish to complain about any aspect of the way in which you have been approached or treated, or how your information is handled during the course of this study, you should contact Dr Susan Hopkins (study email address SIREN@phe.gov.uk) or If you are still unhappy, you can contact the Complaints Manager, Strategy Directorate, Wellington House, 133-155 Waterloo Road, London, SE1 8UG *or email:* complaints@phe.gov.uk What if there is a problem?

**Who is funding the study?**

The Department of Health is funding the study through COVID-19 grant in aid monies to Public Health England. The Chief Investigator is also partially funded through NIHR funding for the Healthcare associated

***Thank you for reading this information and considering taking part.***

# CONSENT FORM: The impact of detectable anti SARS-COV2 antibody on the incidence of COVID-19 in healthcare workers

Short title: SIREN - SARS-COV2 immunity and reinfection evaluation

| **If you agree please initial the box** | |
|---|---|
| I have read the Information Leaflet for Participants | |
| I have been given sufficient time to consider making this decision and have had all my questions answered satisfactorily | |
| I agree to complete a short questionnaire and have a blood sample taken every TWO OR FOUR weeks | |
| I agree to participate in my local hospital PCR screening for the duration of the study | |
| I plan to be available for 1 year | |
| I understand that the information I provide will be shared securely with Public Health England but with no other individual or organisation | |
| I understand that my personal data will be stored in accordance with the Data Protection Act 2018 and the GDPR | |
| I understand that my samples will be tested for the presence of coronavirus and coronavirus antibodies, as appropriate | |
| I have been informed that I can withdraw at any time without giving a reason | |
| I agree for leftover blood to be anonymised and stored at the PHE Seroepidemiology Unit collection for future surveys and assay development | |

Name of participating individual

Date

Signature

Name of person taking consent

Date

Signature

PARTICIPANT ENROLMENT QUESTIONNAIRE

| | **field** |
|---|---|
| Participant number | AANNNNNN |
| Age | NN (years) |
| Gender | M/F |
| Ethnicity | ONS groups |
| Do you have any of the following medical conditions? | Tick boxes for cardiac, respiratory and immunological comorbidities |
| Do you smoke? | Currently / Previously / Never |
| Occupation | drop down list including<br><br>doctor, nurse, HCA, pharmacist, profession allied to medicine (please specify), administrative & clerical, managerial, discharge, porter, other (please specify)<br><br>(free text field for specification) |
| What setting do you spend most of your time in? | drop down list including<br><br>inpatient ward/emergency department/outpatient department/ICU/radiology dept/day ward/ renal dialysis unit/ other (specify) |
| Does your work involve direct contact with patients? | Yes/No |
| Does your work involve regular aerosol-generating procedures at higher risk of acute respiratory infection? | Yes/No |
| Do you consider yourself to be immunocompromised? | Yes/No/Prefer not to say |

| | |
|---|---|
| If yes, please select | Green book categories drop down |
| Have you had laboratory confirmed COVID-19 since February 1st 2020 | Yes/No |
| If yes, date of onset | DDMMYY |
| If no, have you had any illness since February 1st 2020 | Yes/No |
| If yes, date of onset | DDMMYY |
| For either lab confirmed COVID-19 or any other illness, please select all the symptoms which you had during your illness | Fever, cough, difficulty breathing, muscle aches, sore throat, sneezing, runny nose, headache, loss of smell, loss of taste, nausea, vomiting, diarrhoea, tiredness |
| What was your first symptom, if you can recall that | Fever, cough, difficulty breathing, muscle aches, sore throat, sneezing, runny nose, headache, loss of smell, loss of taste, nausea, vomiting, diarrhoea, tiredness |
| If yes, how long where you ill for? | NN (days) |
| If yes, were you admitted to hospital? | yes/no |
| If yes, what is the date that you were tested on? | DDMMYY |
| Where was the test done? | free text |
| Do they think/know they have been infected by SARS-CoV-2? | Yes/ No |

PARTICIPANT QUESTIONNAIRE 2: MONITORING

| | **field** |
|---|---|
| Participant number | |
| Have you had an illness consistent with COVID-19 since your last serology test? (symptoms consistent with COVID-9 are respiratory infection, pneumonia, gastrointestinal infection) | Yes one illness/ Yes more than one illness/ No |
| If yes, symptoms of the most recent illness, tick all that apply | Fever, cough, difficulty breathing, muscle aches, sore throat, sneezing, runny nose, headache, loss of smell, loss of taste, nausea, vomiting, diarrhoea, tiredness |
| If yes, date of onset | DDMMYYYY |
| If yes, how long were you ill for | days |
| Are you still symptomatic | Yes/ No |
| If yes, were you admitted to hospital? | |
| If yes, was the illness laboratory-confirmed as COVID-19? If so, please provide the date and location of the test | |
| If you have had more than one illness, click here (open new list of same fields for illness2) | |
| Have they been in contact with someone that they know (based on a diagnostic test) was infected with Covid-19/SARS-CoV-2? | Yes/No |
| If yes, date of last contact | DDMMYYYY |
| If yes, please tick all that apply | An individual who lives in your own household<br>Someone outside their household |

| | |
|---|---|
| If outside the household, please tick all that apply | Non work contact<br>Work contact outside clinical area<br>Patient without PPE<br>Patient with PPE |

WITHDRAWAL FORM

| Participant number | |
|---|---|
| We understand that you wish to withdraw from the study, please select one option to confirm your withdrawal from the study | • Withdraw from any future study visits, but allow your existing materials and data to remain in the study and be tested, and allow the study team to continue to access your results from your routine PCR testing undertaken by the Trust.<br>• Withdraw from any future study visits, and withdraw the use of your existing materials, but allow the data generated up to the date of withdrawal to remain in the study. Your samples will be destroyed but your record will remain in the study database.<br>• Withdraw from any future study visits, and withdraw the use of your existing materials, and have your data removed from the study. Your samples will be destroyed and your record removed from the study database. |

# Appendix 2: Amendment History

| Amendment No. | Protocol Version No. | Date issued | Author(s) of changes | Details of Changes made |
|---|---|---|---|---|

# References


1. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020;382(8):727-33.
2. World Health Organisation. WHO announces COVID-19 outbreak a pandemic. Available from: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic. WHO: Geneva, 2020 (accessed 11 May 2020).
3. World Health Organisation. Coronavirus disease (COVID-19) Situation Dashboard. Available from: https://covid19.who.int/. WHO: Geneva, 2020 (accessed 11 May 2020).
4. Public Health England. Guidance and standard operating procedure COVID-19 virus testing in NHS laboratories. Available from: https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/guidance-and-sop-covid-19-virus-testing-in-nhs-laboratories-v1.pdf. PHE: London, 2020 (accessed 11 May 2020).
5. Konrad R, Eberle U, Dangel A, et al. Rapid establishment of laboratory diagnostics for the novel coronavirus SARS-CoV-2 in Bavaria, Germany, February 2020. *Euro Surveill* 2020;25:9.
6. To KK, Tsang OT, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020;20(5):565-574.
7. Wikramaratna P, Paton RS, Ghafari M, Lourenco J. Estimating false-negative detection rate of SARS-CoV-2 by RT-PCR. *MedRvix* 2020; https://www.medrxiv.org/content/10.1101/2020.04.05.20053355v2**.**
8. Li Z, Yi Y, Luo X, et al. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. *J Med Virol* 2020;doi:10.1002/jmv.25727
9. Hamre D, Beem M: Virologic studies of acute respiratory disease in young adults. V. Coronavirus 229E infections during six years of surveillance. *Am J Epidemiol.* 1972; 96: 94–106.
10. Treibel TA, Manisty C, Burton M. et al. COVID-19: PCR screening of asymptomatic healthcare workers at London hospital. *Lancet* 2020;18(1):462.
11. de Souza D, Cancho VG, Rodrigues J, Balakrishnan N. Bayesian cure rate models induced by frailty in survival analysis. *Stat Methods Med Res* 2017;26(5):2011-28
12. Christopher H. Jackson. Multi-State Models for Panel Data: The msm Package for R. *J Stat Soft* 2011;38:8.

Public Health England

Public Health Scotland

HSC Public Health Agency

GIG CYMRU NHS WALES | Iechyd Cyhoeddus Cymru Public Health Wales

# SIREN

SARS-CoV2 Immunity & Reinfection Evaluation

# SIREN: The impact of detectable anti SARS-CoV-2 antibody on the incidence of COVID-19 in healthcare workers

## Synopsis

| | |
|---|---|
| Study title | Impact of detectable anti-SARS-CoV-2 on the subsequent incidence of COVID-19 in healthcare workers |
| Short title | SIREN (**S**arscov2 **I**mmunity & **RE**infection Evaluatio**N**) |
| Chief Investigator | Susan Hopkins |
| Sponsor | Public Health England |
| Funder | Department of Health and Social Care |
| Study coordination | Public Health England<br>Public Health Agency (Northern Irish Sites), Public Health Scotland (Scottish Sites), Public Health Wales (Welsh Sites) |
| Statistics and design | Andre Charlett |

| | |
|---|---|
| Study management group | Study design and management: Susan Hopkins, Meera Chand, Colin Brown, Victoria Hall<br>Statistics: Andre Charlett<br>Epidemiology: Mary Ramsay, Victoria Hall<br>Virology and serology: Maria Zambon, Tim Brooks<br>Data management: Anne Marie O'Connell<br>Genomics: Richard Myers<br>Information governance: Samantha Organ<br><br>Devolved Administrations: Muhammad Sartaj (Northern Ireland), Josie Murray (Scotland), Eleri Davies (Wales) |
| Study Design | Prospective cohort study to determine incidence of new SARS-CoV-2 infection in seronegative and seropositive staff working in healthcare organisations |
| Study participants | Staff working in healthcare organisations in the UK |
| Sample size | 100,000 staff working in healthcare organisations |
| Planned study period | 24 months |
| Planned recruitment period | 12 months for each participant from enrolment; interval analysis of incidence and prevalence monthly and statistical analysis on primary outcome 3 monthly. |
| Study Protocol version | 6.2 |

# Contents

# Summary

COVID-19 is causing a global pandemic. This study aims to find out whether staff working in healthcare organisations who have evidence of prior COVID-19, detected by antibody assays (positive antibody tests), are protected from future episodes of infection compared to those who do not have evidence of prior infection (negative antibody tests). With the introduction of COVID-19 vaccinations for healthcare workers from December 2020, this study will also examine immunity acquired by a vaccine and obtain early estimates of vaccine effectiveness. It will explore both short and long-term effectiveness of a vaccine against infection and immunological response to a vaccine, including potential differences in response associated with factors such as prior exposure and antibody status.

In this study, we will recruit staff working in healthcare organisations who will be followed for a year and study their immune response to the virus causing COVID-19, called Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2). We will do this by collecting data on their history of COVID-19 infection and any new symptoms. All participants enrolled into the study are being asked to have a regular nose swab or a combined nose and throat swab (depending on local decision) every other week to detect mild cases or cases who do not have symptoms. This is the main test that is currently used to detect and diagnose infection by looking directly for the virus in the nose and throat. Once the infection is cleared, we cannot detect the SARS-CoV-2 virus in samples. Therefore, we will also ask these individuals to have blood samples taken every other week to determine whether they have antibodies to the infection. The frequency of swab tests and/or blood tests may be changed to between 1 and 4 weekly tests. This will depend on national and local epidemiology, study retention, feedback and results. These blood samples allow the previous infection to be detected as the response to infection in the body is to produce small particles in the blood called “antibodies”. It takes up to 4 weeks to make enough antibodies to fight the infection. But once someone recovers, antibodies stay in the blood at low levels and this may help prevent us from getting infected with the same infection again. However, for SARS-CoV-2 infection, we do not know yet if the detection of specific antibodies is a correlate of protection from future infection. Through this study we will provide this very important information which will help to understand the future impact of COVID-19 on the population.

The first SIREN participants were recruited on 18 June 2020 and therefore end their one-year follow up in June 2021. Insights from SIREN have proved vital for informing Government policy and have made significant contribution to the understanding of immunity to SARS CoV-2. The national and international need for this data remains high, particularly with the emergence of new variants and the need to understand the durability of immunity following infection and vaccination. Participants will therefore be offered an opportunity to remain in SIREN for up to an additional 12 months, upon completion of their original follow-up period to maintain and strengthen the evidence base provided by SIREN.

# 1. Background

On 31st December 2019, the first cases of infection with a novel coronavirus, subsequently designated SARS-CoV-2, emerged in Wuhan, China.[1] A global pandemic was declared by the

World Health Organisation (WHO) on 11th March 2020.[2] By 9th May, there were more than 3.6 million confirmed infections and 240,000 reported deaths globally.[3]

In order to determine the population impact of potential future waves of the pandemic of COVID-19, there is an urgent need to understand whether prior infection and any host immune response provides protection from future reinfection with SARS-CoV-2.

Currently, a confirmed diagnosis of infection of SARS-CoV-2, relies on laboratory diagnosis of infection using reverse transcriptase polymerase chain reaction (PCR). Diagnostic PCR typically targets the viral ribonucleic acid (RNA)-dependent RNA polymerase (RdRp) or nucleocapsid (N) genes using swabs collected from the upper respiratory tract (nose and throat).[4, 5] However as PCR only detects acute infection we are unaware of the true population prevalence of infection. While there is increasing evidence that a proportion of the population can be infected and asymptomatic, the true proportion of asymptomatic infection in the population is unknown.

Across the UK, some NHS organisations are choosing to screen healthcare workers with RT-PCR to identify individuals who are positive for SARS-CoV-2 and reduce the risk of healthcare worker (HCW) transmission to other HCW and patients through identification and appropriate exclusion from work. However, PCR from upper respiratory tract swabs may be falsely negative, due to quality or timing of collection; or due to viral titres in upper respiratory tract secretions peaking in the first week of symptoms,[6] but declining below the limit of detection in patients who present with symptoms beyond this time frame.[7] In individuals who have been infected and recovered, PCR provides no information about prior exposure or immunity. In contrast, assays that reliably detect antibody responses specific to SARS-CoV-2 could contribute both to diagnosis of acute infection (via rises in IgM and IgG levels) and to identify those who have been exposed and recovered with or without symptoms (via persisting IgG).[8,9] A further research gap is whether the detection of an antibody response to SARS-CoV-2 will provide immunity to re-infection. In a study of medical students, reinfections with HCoV-229E were detected in 937 student years of follow-up; neutralizing antibody to 229E was common but it did not appear to influence the occurrence of, or likelihood of illness with, reinfection as judged by complement fixation seroconversion.[10] While IgG and IgM to SARS-CoV-2 can be effectively used to determine past exposure, it may be that these antibodies do not protect against future infection. The presence of neutralising antibodies may be required as these not only confirm prior infection but also prevent future infection. There are no commercial assays for live virus neutralisation assays at present but PHE and other university partners are developing in house assays to detect these specific antibodies. We will examine correlates of different types of antibody response and the protective effect against re-infection.

Therefore, this study will answer the key question on whether prior SARS-CoV-2 infection confers future immunity to SARS-CoV-2 re-infection in staff working in healthcare organisations who have a higher prevalence of infection than the general population; in one study 20% of asymptomatic London HCW tested for SARS-CoV-2 were positive on PCR tests over a 5-week period from 23rd March to 3rd May, with incidence peaking at 7.1% and declining 6 fold after 5 weeks.[11] It will also determine whether reinfection is due to a lack of neutralising antibody or waning immunity over time in individuals who are anti-SARS-CoV-2 positive, whether such individuals can be re-infected symptomatically or asymptomatically, about the incidence of new infection in individuals without

prior exposure and the severity of COVID-19 in initial and subsequent episodes of infection if re-infection occurs. We will also use genomics to determine the viral diversity in staff working in healthcare organisations at each sampling point and particularly if individuals are re-infected.

Healthcare workers (HCW) have been identified as one of the first groups targeted for vaccination and therefore will be an important group within which to obtain early estimates for vaccine effectiveness[12]. They also may be a group whose vaccination details may not be easily identified in other data sources (such as GP data). Therefore, from December 2020, the scope of this study will include examination of immunity and re-infection associated with healthcare workers who have received a COVID-19 vaccine.

On 18th June 2021 the first SIREN participants reach the end of their 12-month follow up period. As of May 2021, important questions about COVID-19 immunity remain, particularly with regards to the longevity of immune response from both natural infection and vaccination. New questions continue to emerge, particularly around the strength of immunity against new variants and the efficacy of vaccination against these strains. In order to ensure SIREN continues to provide critical insights into immunity to SARS-CoV-2, participants will be offered the opportunity to take part in an additional follow up period of up to 12 months. Follow-up testing will continue to collect regular PCR and serology samples with methods and frequency subject to local site variation.

# 2. Aim and Objectives

**Aim:** The overall aim of this study is to determine if prior SARS-CoV-2 infection in health care workers confers future immunity to re-infection.

**Objectives:**

**Primary:** To determine whether the presence of antibody to SARS-CoV-2 (anti-SARS-CoV-2) is associated with a reduction in the subsequent risk of re-infection over short term periods (reviewed monthly), the next year and in the longer-term.

**Secondary:**

1. To estimate the prevalence of SARS-CoV-2 infection in staff working in healthcare organisations by region, using baseline serological testing at study entry and symptom history from January 1st 2020 to date of study entry
2. To estimate the subsequent incidence of symptomatic and asymptomatic SARS-CoV-2 infection and determine how this varies over time, using regular PCR testing (combined with any intercurrent symptomatic testing)
3. To estimate cumulative incidence of new infections in staff working in healthcare organisations stratified by age, sex, staff group, ethnicity and co-morbidities.
4. To measure the ability to culture viable virus from cases of re-infection diagnosed by PCR and whether those that are persistently positive on PCR are continuing to shed viable virus
5. To use genomic comparison to determine whether healthcare workers who become PCR-positive for a second time within a defined time frame are experiencing persistent infection or re-infection
6. To determine how serological response changes over time
7. To determine whether there is a relationship between serological response (using enzyme immunoassay detection of IgG) and the presence of neutralising (protective) antibodies
8. To identify serological, demographic or clinical factors that correlate with the presence of neutralising antibodies, including subsequent disease severity
9. To investigate the phylogenetic relatedness of SARS-CoV-2 viruses causing staff working in healthcare organisations infections
10. To monitor effectiveness of a vaccine/vaccines against an infection and symptomatic disease
11. To monitor immune response to vaccination over time

# 3. Methods

## 3.1 Study design

The overall study design is a prospective longitudinal cohort study of staff working in healthcare organisations within the UK.

## 3.2 Study sites

Any NHS organisation in the UK (including all four nations) who can deliver the requirements of the study will be eligible to participate. Primary care sites (GP surgeries or clinical commissioning groups (CCGs)) as well as independent healthcare providers, will be eligible to participate if they can demonstrate the ability to deliver the study.

Each site will be asked to invite eligible staff to participate. While initially we are not providing quotas, we may review the cohort of participants recruited initially to ensure appropriate representation by staff group, age, sex and ethnicity; quotas may be provided at that stage if certain groups are under-represented.

Organisations which fulfil the criteria of the sites described above will also be able to become Participant Identification Centres (PICs), where they will publicise the study to staff (using the methods outlined in this protocol), identify potential eligible participants and refer them to a study site. In order to be a PIC site the organisation will be expected to attend a Site Initiation Visit and complete a PIC agreement with the referral study site.

## 3.3 Inclusion/exclusion criteria

Inclusion criteria:

- **Healthcare organisation staff member who works in a clinical setting where patients are present**
  This includes all staff members who work on a site where patients are present – they do not need to be patient facing. They must work in the healthcare organisation but do not need to be employed by that organisation (working student healthcare workers or those employed by contracting organisations are eligible).
- **Can provide written consent** (this is collected in an online survey)
- **Is willing to remain engaged with follow-up for one year, even if they move to another healthcare organisation**

Exclusion criteria*

- **Written consent not provided**
- **Temporary short-term staff member**

For study purposes this would be a staff member unlikely to remain at the healthcare organisation for at least three months, unless they are expected to transfer to another healthcare organisation; bank, agency and locum staff are eligible as long as they fulfill this criterion.

*While SIREN is happy to accept participants who are enrolled in other studies, individuals who are already enrolled in studies may not be able to participate if the other study does not accept co-enrolment into SIREN.

SIREN is collaborating with a number of wider immunology studies which are conducting further investigation into immunity response of SARS-CoV-2. SIREN participants may be invited to join one or more of these studies, in line with the relevant eligibility criteria. Data may be shared between SIREN and associated studies, in line with information governance and data protection procedures outlined in this protocol. A full list of associated studies can be found in Appendix 1.

## 3.4 Sample size

A simulation approach has been used to estimate the power to detect relative differences between the study cohorts.

It is anticipated that at least 115 NHS organisations will join this study. We are aiming for up to 100,000 staff working in healthcare organisations to be recruited. We have estimated that 25% of our cohort will be seropositive at enrolment healthcare (based on 20% of staff who were asymptomatic and tested positive in one London hospital between 23rd March and 2nd May 2020[11]).

For the longer-term outcomes of the study, the initial sample sizes were calculated on the basis of recruiting 10,000 participants from 40 sites. It has been assumed that a minimum of 250 participants would be recruited from each selected healthcare organisation, with a standard deviation of 50. The proportion of seropositive recruits at each site has been obtained from a Gaussian distribution with a mean of 0.25 and standard deviation of 0.05 to reflect expected inter-site variation. The cumulative incidence in each Site in the seronegative cohort has been simulated using Gaussian distributions with means of 0.05, 0.1, 0.2 and 0.3 each with a coefficient of variation of 0.2. This range represents that which is feasible to observe over a 12-month period, given the behavioural and social interventions still being employed during the study to control transmission.

A study duration of 52 weeks has been assumed with the inter-test period of 2 weeks. A total attrition of 35% of participants is assumed, unaffected by sero-status and occurring at a constant rate over the 52 weeks of the study. Attrition has been assumed to be independent of the infection process.

A range of cumulative incidence in the seronegative cohort has been used to reflect the immune effectiveness; with units in the simulations being allocated to be infected or not, using a draw from a Bernoulli distribution with p equal to the Site and cohort specific simulated cumulative infection rate. A simplifying assumption of a constant infection rate over the study period has been used.

For each scenario a set of 200 simulations were performed. For each simulation, the total number if infections and person weeks of follow up was calculated for each cohort in each organisation.

This data was analysed using a mixed effects Poisson model, using the natural logarithm of the person weeks as an offset. Power was estimated as the proportion of simulations for which the Wald statistic p value for the estimated incidence rate ratio of the seropositive to seronegative cohorts was less than 0.05. These are presented in Table 1, indicating that there is sufficient power for all but the smallest immune efficacy of 0.1 i.e. a 10% reduction in incidence in the seropositive cohort. Such a small reduction is indicative of an irrelevant level of protection to provide a means of controlling the pandemic via natural herd immunity.

**Table 1: Power estimates obtained via simulation for a range of immune effectiveness and cumulative incidence**

| Cumulative incidence in the seronegative at baseline cohort (per 100 participants) in 12 months | Immune Effectiveness 10% | Immune Effectiveness 20% | Immune Effectiveness 30% | Immune Effectiveness 40% | Immune Effectiveness 50% |
|---|---|---|---|---|---|
| 0.05 | 0.15 | 0.44 | 0.79 | 0.98 | 1.00 |
| 0.1 | 0.20 | 0.77 | 0.99 | 1.00 | 1.00 |
| 0.2 | 0.53 | 0.99 | 1.00 | 1.00 | 1.00 |
| 0.3 | 0.67 | 1.00 | 1.00 | 1.00 | 1.00 |

To investigate the ability of a study of this size to detect differential immune efficacy in subgroups of healthcare workers, for example those aged over 50 or from a BAME group, a further set of simulations were undertaken. It has been assumed that the subgroup of interest is on average 35% of the total cohort size. The cumulative incidence used in the previous simulations has again been used with just those in the subgroup that are seropositive having an absolute reduction in immune efficacy of 5%, for example, if the cumulative incidence in the seronegative cohort is 0.3 (30%) and the immune efficacy in the non-sub group is 0.5 with a cumulative incidence of 0.15 (15%) in this group, the cumulative incidence in the seropositive sub group is 0.2 (0.15+0.05 20%). The interaction between cohort type and subgroup has additionally been added to the model and the estimated power obtained using the proportion of the 200 simulations for which this interaction had a Wald test p value of less than 0.05. Only immune efficacies of 0.3 or greater have been used, and these are presented in Table 2.

**Table 2: Power estimates obtained via simulation for a range of immune effectiveness and cumulative incidence for detecting a sero-status by subgroup interaction**

| Cumulative incidence in the seronegative at baseline cohort (per 100 participants) in 12 months | Immune effectiveness 30% | Immune effectiveness 40% | Immune effectiveness 50% |
|---|---|---|---|
| 0.1 | 0.93 | 0.91 | 0.98 |
| 0.2 | 0.70 | 0.82 | 0.84 |
| 0.3 | 0.67 | 0.68 | 0.69 |

In order to determine the outcome of immunity in much shorter intervals, by increasing the cohort recruitment to 100,000, we will be able to detect a difference between 0.05% and 0.1% cumulative incidence; even taking the incidence to as low as 0.02% in the seronegative group there is still excellent power of around 94%.

Estimated power for a two-sample proportions test

Pearson's chi-squared test

Ho: p2 = p1 versus Ha: p2 != p1

| alpha | power | N | N1 | N2 | nratio | delta | p1 | p2 |
|---|---|---|---|---|---|---|---|---|
| .05 | .9403 | 1.0e+05 | 25000 | 75000 | 3 | .001 | .001 | .002 |
| .05 | 1 | 1.0e+05 | 25000 | 75000 | 3 | .002 | .001 | .003 |
| .05 | 1 | 1.0e+05 | 25000 | 75000 | 3 | .003 | .001 | .004 |
| .05 | 1 | 1.0e+05 | 25000 | 75000 | 3 | .004 | .001 | .005 |

If the incidence is lower than this, i.e. 0.05% in the seropositive group and 0.1% in the seronegative the power becomes sub-optimal at around 66%, but is sufficient for large differences.

Estimated power for a two-sample proportions test

Pearson's chi-squared test

Ho: p2 = p1 versus Ha: p2 != p1

| alpha | power | N | N1 | N2 | nratio | delta | p1 | p2 |
|---|---|---|---|---|---|---|---|---|
| .05 | .663 | 1.0e+05 | 25000 | 75000 | 3 | .0005 | .0005 | .001 |
| .05 | .9933 | 1.0e+05 | 25000 | 75000 | 3 | .001 | .0005 | .0015 |
| .05 | 1 | 1.0e+05 | 25000 | 75000 | 3 | .0015 | .0005 | .002 |
| .05 | 1 | 1.0e+05 | 25000 | 75000 | 3 | .002 | .0005 | .0025 |

Estimates for the vaccine effectiveness element are based on an assumed population of 40,000 participants.

**Table 3 – Precision estimates assessing 95% CI around a vaccine effectiveness (VE) of 60% and 90%**

| Incidence in unvaccinated | Cases in unvaccinated | 95% CI around VE of 60% | 95% CI around VE of 90% |
|---|---|---|---|

| | | | |
|---|---|---|---|
| **0.5** | **32** | **39-74** | **81-95** |
| **1** | **65** | **46-70** | **85-93** |
| **2** | **130** | **50-68** | **86-93** |
| **5** | **325** | **54-65** | **88-92** |

95% CIs will become narrower as VE increases (wider as it decreases) and also wider if coverage increases and in any strata. Overall the table shows that reasonable precision should be achievable.

For the initial vaccine effectiveness estimate three months post vaccine, the focus is on those seronegative at baseline.

Estimates for vaccine effectiveness are based on the following assumptions; 65% are seronegative at baseline, based on the baseline of 70% and assuming an additional 5% since this time. 75% are vaccinated and incidence during the 3 months in unvaccinated is 0.5%, 1%, 2%, 5%. This is based on incidence seen in September 2020 of 0.25% per 2 week (0.5% in a month or 1.5% in 3 month) to the incidence of 0.85% in 2 weeks in October (1.7% in a month or 5% in 3 months).

N=40,000 (26000 seronegative of whom 29,500 are vaccinated and 6,500 unvaccinated) is assumed.

## 3.5 Site set up

A site initiation visit (SIV) will take place by conference call for each study site. During the SIV, a member of the central study team will describe the study methods and discuss possible ways to organise recruitment and data collection with the local clinical and research team. An electronic Recruitment Log (a participant linelist on a formatted MS Excel spreadsheet) will be provided to all sites to record details on all recruited participants. The recruitment log will be pre-populated with study number and passcode**.** Both fields are unique codes required by participants to access and complete the online consent form and enrolment questionnaire. Study sites will be responsible for updating and securely storing this information in the recruitment log and for securely transferring it to PHE SIREN team and/or the appropriate Devolved Administration public health organisation's SIREN team securely if requested.

## 3.6 Recruitment

All recruitment should take place as quickly as possible, although Public Health England have confirmed an extension of recruitment to 31st March 2021. Participants are recruited by an all-staff communication requesting volunteers, but sites may choose to use locally appropriate publicity methods, e.g. posters, social media, intranet, stalls etc. Local organisations can produce their own publicity materials, although PHE have developed approved posters for use.

The organisation's research team will be responsible for assessing eligibility and ensuring participants are informed about the study requirements, ensuring all participants receive a participant information leaflet prior to recruitment. The research team should then provide

interested and eligible staff with a unique study number, passcode and the webpage to complete the online consent form and enrolment questionnaire, and their details should be recorded in the Recruitment Log. Participants will not be able to access the online survey without these two fields, which can only be provided by their local research team. Care should be taken by the local organisation's research team to ensure that the unique study number and passcode combination are not issued to multiple potential participants, and if once given to one individual should never be re-allocated (even if never used).

Following the introduction of a vaccine, all vaccinated healthcare workers should receive a leaflet informing them about SIREN and that someone from their local research team may be in contact to discuss enrolment. If people wish to opt out of this contact they can do so via local processes set up at each site, e.g. email.

**Screening, eligibility assessment and formation of the cohort at each hospital:**
For each participant, a member of the healthcare organisation's research team will assess eligibility and monitor the proportion of participants from different occupational groups; they may choose to send out further communications to support recruitment to under-represented groups. The healthcare organisation may also choose to enrich their seropositive participants at baseline, particularly if they have identified low seroprevalence among staff working in their healthcare organisation. This may be done by general communication encouraging staff who know they are seropositive to enrol, or by targeted communication to known seropositive staff, if this data is held and accessible according to local organisational protocols. The study is observational only and there is no randomisation or allocation to any subgroups.

Staff such as junior doctors may have planned moves to other healthcare organisations within the study period. Where the participant is moving to a new organisation participating in SIREN, it is the responsibility of the original enrolling organisation's research teams to transfer these participants (including PII) directly to the new sites and communicate this transfer to the SIREN team. The participant should keep their original study number for the study duration. Where staff may be moving to a healthcare setting not participating in SIREN, they can remain enrolled, subject to the approval of the principal investigator and their original enrolling organisation continuing to offer them testing. Frequency of testing for these participants can be flexible. Participants who are leaving healthcare practice (e.g. on retirement) can also remain enrolled in the study, subject to the approval of the site's principal investigator.

## 3.7 Enrolment and consent

Participants will be enrolled in SIREN once they have submitted the online consent form and enrolment questionnaire, which requires them to provide their contact details for follow-up communication.

Consent will be undertaken online as the first part of the enrolment questionnaire, with participant initials requested against each point. A link to a copy of the consent form will be automatically sent to both the participant and study team on submission of the enrolment consent form and questionnaire.

The local organisation's SIREN study team will check the validity of the responses of the consent form on receipt. If a participant has provided responses 'Yes' or 'Y', or 'x' throughout, this can be accepted as valid consent. In the event that the participant has given a response which is ambiguous (e.g. 'N', 'No' or '?'), the local study team will contact the participant, clarify whether they intended to consent and obtain a paper copy of the consent form or facilitate withdrawal. They will send a secure email to the central PHE study team containing participant study number and the steps taken to ensure valid consent (they do not need to transfer the consent to PHE).

Where the PHE enrolment questionnaire is not functioning (e.g. server down or internet not working), the research team may obtain consent using a paper copy of the consent form and proceed to obtain bloods, but must follow up and ensure that the participant has completed the enrolment questionnaire within 48 hours. This paper consent form does not need to be sent to the PHE SIREN team but should be retained securely at the site. A paper copy of the consent form is only to be used in exceptional circumstances as outlined above.

A supplementary consent form has been designed for participants previously recruited into PHE seroprevalence surveys, as detailed in section 6.

At the end of their 12-month follow-up period, active participants may be given the option to participate in an additional follow-up period of up to 12 months. Explicit consent will be obtained for one of three options; to end all involvement with SIREN, to continue with active follow-up or to continue with passive follow-up only. Passive follow-up will involve no further collection of SIREN samples but permits SIREN to link existing study data, including to potential future test results not collected as part of SIREN. This consent will be collected electronically via Snapsurvey and is not conditional to a participant ending their active involvement with SIREN. Where possible, participants will be notified of this choice in advance of the end of their original follow-up period. In the event of the online questionnaire being unavailable, a paper copy of consent to extended follow-up may be collected via sites following the process for consent outlined above.

## 3.8 Data collection and measurements

### 3.8.1 Baseline data

**Self-completed enrolment questionnaire**: The participant will self-complete an online survey at enrolment. Baseline data includes demographic/social factors (gender; date of birth; ethnicity; smoking history); comorbidities including whether immunocompromised; job role; potential exposures to COVID-19; history of COVID-19 symptoms and testing since January 2020.

**Baseline testing**: Upon receipt of the consent form, the local organisation's research team will organise for the participant to attend appointment(s) to provide blood samples and swabs for viral testing. Ideally these samples will be provided within 48 hours of the participant's completion of the enrolment questionnaire; up to 5 days after enrolment will be accepted as an outer limit. If a participant has submitted samples for an antibody test up to 5 days prior to enrolment, this can count as their baseline antibody test, as long as there is sufficient serum available to be sent to PHE. Any samples for baseline testing taken over 5 days following enrolment should be recorded as a protocol deviation. Where possible, a baseline sample should be collected prior to receiving a vaccine, however a vaccine should not be delayed to accommodate this.

**SARS-CoV-2 molecular screening**: This will be undertaken according to the local organisation procedure for screening healthcare workers or through a protocol set up specific for this study. Screening will be performed at regular intervals from 1 to 4 weeks (as specified by the PHE SIREN study team) and at any point the participant has clinical symptoms that meet local criteria for testing. It may be a nose swab or a nose and throat swab and may be undertaken by self-swabbing or administered by a trained member of staff depending on the local organisational preference. The swab will be tested for SARS-CoV-2 RNA in an accredited laboratory (or those who have provided satisfactory verification documents), which may be the clinical laboratory or through the national testing programme, using a validated nucleic-acid amplification test (NAAT), such as RT-PCR. As this is a clinical sample that requires reporting of the result individually, full standard identifiers will be used. Residual specimen and any available nucleic RNA extracts should be stored at -80$^{O}$C.

Whole viral genome sequencing should be performed on SARS-CoV-2 positive specimens through the site's arrangements with their local COVID-19 Genomics UK Consortium (COG-UK) centre. The SIREN central study team can link sites to their local COG-UK centre if needed.

**Anti-SARS-CoV-2 antibody testing**: Each participant will have a venous blood test on the same day as the PCR screening (within 48 hours is acceptable). The local SIREN team will determine the optimal procedure within each healthcare organisation, which may be visiting a study nurse, attending the phlebotomy service with a pre-designed form, or ward teams/peer-to-peer blood testing is also acceptable if this is approved by local governance procedures and those taking blood have been appropriately trained. Up to 10ml venous blood sample will be taken (in one or more blood tubes) and serum will be separated. For participants in wider immunology studies, the volume of some blood draws will be greater, up to 100ml (typically 50-60ml). No more than 470ml blood will be taken in a 16-week rolling period. The volume required for local testing will be removed and the residual volume should be retained in the form of one full 2ml aliquot for use as below. The aliquot for local testing will be tested immediately in the organisation's local laboratory using an immunoassay. As this is a clinical sample that requires reporting, full standard identifiers will be used.

The second aliquot will be stored for a maximum of 7 days at 2-8$^{O}$C, or at -20$^{O}$C or below for longer, and batch shipped to PHE for later analysis using alternative assays to establish correlates of protection and compare the serological response results across different assays. This will be completed initially for 10,000 participants to validate the assays, and the composite PHE serological assessment will be used to assign these participants to the seronegative or seropositive cohort at baseline. The results of this validation on 10,000 participants will inform whether future participants also receive composite PHE serological assessment or are assigned a serological status based on their local results (potentially adjusted for assay-specific validation results).

This additional PHE serological evaluation will be stopped or adjusted when the study team are of the view that there is sufficient data to predict the relationship of the commercial assay result to the serological characterisation. A small seropositive sub-cohort will be selected to follow for a detailed assessment of antibody dynamics.

Participants may be offered additional testing via the NHS Test and Trace testing platform using a finger prick collection method, if the test is not offered as standard by their trust. This optional test would follow procedures as agreed with the national testing platform and by offered in addition to the venous blood sampling described above.

3.8.2 Follow-up data

Participants will be asked to complete follow-up surveys and provide swabs and blood samples for repeat testing at regular intervals.

**Follow up questionnaire**: The participant will receive a unique link sent to their mobile phone or email address to complete their follow-up questionnaire. Initially the follow-up questionnaire will be sent every two weeks. The follow-up questionnaire will collect information on possible COVID-19 symptoms and exposures, risk factors, enrolment in COVID-19 vaccine/treatment trials and any COVID-19 or seasonal influenza vaccine during the follow-up interval.

**Frequency of follow up testing visits:** At the initiation of the study this will occur every two weeks for every site. The frequency may be altered to be more or less frequent (1 - 4 week intervals). This will depend on national and local epidemiology, study retention, feedback and results. This decision will be reviewed by the central PHE SIREN Study Management Group and, if and when, a decision is made, will be communicated to sites. Follow-up visits should not occur more than two days before the scheduled date or more than five days afterwards. If participants are unable to attend for testing within this period (e.g. annual leave), this should be recorded as a missed visit. For their next testing appointment, they can either be started on a new 14-day test cycle or returned their old one. Unless there is a clinical reason to do so, participants should not be tested more than once in a 7-day period for the purpose of SIREN.

**Follow-up PCR testing:** Procedures for follow-up SARS-CoV-2 RT-PCR (or equivalent NAAT) is the same as baseline.

Viral culture may be attempted for those individuals who are persistently NAAT positive, have potentially been re-infected or have been infected since receiving a vaccine; this may involve PHE contacting local research teams and requesting extra swab samples from these participants if there is no suitable residual sample available. Where this is required the participant may need to complete a test at home (as may be self-isolating). We would also request local organisation research teams to set-up procedures with their laboratory to identify these events of interest and report them to PHE SIREN team to ensure timely follow-up. Extra swab tests may also be requested for confirmatory or further testing.

Healthcare organisation research teams are expected to actively monitor testing attendance and test results in their cohort, to follow-up repeat non-attenders, and to alert the PHE SIREN team to events of interest (i.e. a possible re-infection or persistent infection). This may involve working with Virology/Microbiology and Occupational Health, depending on local procedures. If

a participant misses eight consecutive scheduled testing visits or four months of visits (whichever is shorter) this should be recorded as a protocol deviation.

**Follow-up serology**

All samples should be tested locally for antibodies. Whilst sites are requested to retain baseline sera from all participants, at follow-up, sites are only asked to retain and send to Public Health England the sera from participants:

- (First samples for all participants)
- Who are seropositive;
- Who have previously been seropositive (including prior to the study commencing if known);
- Who have been PCR positive at any point (including prior to the study commencing if known);
- Any participant who has been vaccinated against SAR-COV2 should also have their serum retained from the point of vaccination onwards.
- Participants where PHE SIREN team has contacted the local healthcare organisation directly to request samples

The laboratory manual includes details on how to manage this process.

Samples should be shipped in monthly (or more frequent if required) batches to PHE for further characterisation (see laboratory manual for directions).

**Extended follow-up**

The extended follow-up period will continue to collect study data as outlined in the main follow-up, i.e. follow-up survey, PCR testing and serology testing. These may be at reduced frequencies or via different collection methods. Specific testing schedules and procedures may differ by site and will be dependent on local arrangements at each trust, as outlined in the site manual. Serological follow-up may be at a reduced frequency compared to the main follow-up period but should occur at a minimum of every three months. PCR testing should occur at a minimum of monthly and may continue to take place in a supervised environment at sites, or via national postal testing offered by NHS Test and Trace.

## 3.9 Reporting of results

The local PCR and serological results will be reported according to the local healthcare organisation's standard occupational health procedures and the methods of communication should be agreed and disseminate by the research team. Requesting tests and accessing results may require an appropriate team (according to the organisations processes this may be the research team and/or occupational health team and/or another designated clinical team) to have access to the individual’s electronic health record/pathology system.

**Action on positive results**: If the PCR screen is positive, the participant will be advised by the local healthcare organisation and/or NHS Test and Trace (or the relevant contact tracing service in the devolved administrations) of the result and to follow the national guidance for self-isolation and the organisation guidance for return to work, with contact tracing and action as required. If the serology assay is positive, the participant will receive advice from the local healthcare organisation indicating that this does not mean they are immune to reinfection, and that they must continue to adhere to infection control measures at home and work as usual. This may involve the healthcare organisation's research team, Occupational Health, or Laboratory/Virology teams, dependent on local organisational procedures. As previously described, positive specimens/nucleic extracts should be sent for sequencing and viral culture may be attempted by PHE for those individuals who are persistently NAAT positive, or have potentially been re-infected.

## 3.10 Participant withdrawal

Participants may withdraw at any time and this is explained in the participant information leaflet. Participants are directed to complete an online survey accessed from the SIREN webpage and advised their withdrawal will only be finalised once they have completed all the questions. On a request to withdraw they will be given the following options:

1. Do you wish to withdraw from receiving any future follow-up questionnaire requests and serology testing? Yes/No
2. Are you happy for us to undertake future testing of your existing samples? Yes/No
3. Are you happy for the study team to continue to access your results from any future routine PCR tests undertaken by your healthcare organisation? Yes/No
4. Are you happy for all of your survey responses and linked laboratory data generated up to the date of withdrawal to remain in the study? Yes/No

For the extended follow-up period, participants will be given an explicit option to continue their involvement with SIREN, as either active or passive follow up. Participants will also have the option to end their involvement at the end of their original 12-month follow up period. Among participants opting to take part in extended follow-up, their rights over data and withdrawal for the main study component will be unaffected. If participants wish to withdraw from the extended follow-up they may do so at any time, following the withdrawal process and options outlined above.

## 3.11 Storage of materials and additional Laboratory testing

All residual serum, positive swab material, and nucleic acid extract as outlined in the previous sections will be stored at the hospital site and shipped to PHE or agreed sequencing laboratories in batches. They will be used for confirmatory testing and serological characterisation, or for viral genome sequencing and/or culture. No human DNA genomic investigations will be undertaken.

**Additional serological testing:** In addition to the baseline cohort serological characterisation described above, seropositive participants in whom reinfection is identified, plus a cohort of matched non-infected seropositive controls, will have their sera further characterised using additional assays and for the presence of neutralising antibody, to provide hypothesis generating

data on mechanisms of protective immunity. Enhanced biochemical and functional serological testing will also be conducted in a subset of participants through associated studies (Appendix 1).

**Genomic analysis**: All positive samples from participants will be sequenced as part of the routine sequencing of NHS residual samples in COG-UK Consortium laboratories. For participants who have more than one positive PCR test, genomes will be compared where possible to provide evidence to support reinfection or persistent infection. Phylogenetic analysis of SARS-CoV-2 from staff in healthcare organisations, using the study samples and the wider collection of genomes available through the COG-UK Consortium, will also be undertaken as an exploratory analysis into the diversity and spread of SARS-CoV-2 in healthcare workers.

**Viral Culture:** Participants with possible re-infection or persistent infection will be identified and viral culture requested. This may be on residual sample from the swab already taken, but in certain circumstances we may request another swab is taken and sample sent for culture if viral culture is not possible due to the type of sample.

**T-cell assays and other studies:** Those individuals who are persistently NAAT (nucleic acid amplification test) positive, or have potentially been re-infected, or have discordant serology may be contacted by the SIREN Study Team to link into optional associated studies (Appendix 1) e.g. assessing T cell responses, B cell responses or further antibody characteristics

### 3.12 Retention

Participants will be involved with SIREN for 12 months, or up to 24 months for those involved in the extended follow-up. During this time, strategies will be employed to ensure sites and participants remain informed, feel valued and their contribution is recognised.

Neither study sites nor participants will be offered an incentive payment for their participation in SIREN. However, the contribution of sites and participants may be recognised with tokens of appreciation including, but not limited to, stickers, badges, lanyards or certificates. Tokens should not be of sufficient financial value to be considered an incentive payment and should not offer benefits to participants that may constitute unreasonably favourable treatment in the workplace. These may be offered by the central PHE SIREN team, or by sites at the discretion of the local SIREN study team.

# 4 Data management

All study documentation will be stored at each site, either in hard copy in a secure environment and/or in electronic copy in an access-limited location on a healthcare organisation server, as decided suitable by the organisation. Use of online platforms for storing study documentation, which should not include PII, could be considered, subject to the healthcare organisation research team providing PHE sufficient reassurance that appropriate data security safeguards are in place. Information received by PHE as the central study site will be stored securely in access-limited locations on PHE servers. Healthcare organisations are required to complete and sign an Organisation Information Document before they are approved to begin recruiting participants into SIREN. This includes a data processing agreement and a data sharing agreement which stipulate requirements around data confidentiality and research governance. PHE, as the sponsor, has the right to audit study site compliance with the terms of the agreement.

Study data is as follows:

- Recruitment log, containing personal and demographic information, retained by the healthcare organisation research team, and transfer may be requested by PHE SIREN team. Information field to collect are NHS number, name, date of birth, sex, date of enrolment and staff type. The local organisation's SIREN team may also choose to optionally collect ethnicity in order to monitor for quotas or to monitor cohort demographics.
- NHS and Lighthouse laboratory records, identified by personal information – this data (including name, date of birth and postcode) is automatically transmitted to PHE as part of the official notification of infection (under Health Protection Regulations) through the Second Generation Surveillance System (SGSS) – it includes both results of PCR tests and serology results performed by the laboratories. The study will access the routine data generated from the national screening programme in the participant organisations with no additional laboratory requirements. Ideally, these data collected by PHE will include the sample type, assay used, cycle threshold (CT) value and optical density as appropriate for each sample tested. For Devolved Administrations (Scotland, Northern Ireland and Wales), laboratory data on healthcare organisation testing may be organised either via: organising reporting from Devolved Administration laboratories through to SGSS, or the Public Health Agency undertaking the laboratory matching for their participants, and then providing linked data back to the PHE SIREN team. Details of data processing and approvals for this will be agreed and documented with respective agencies.
- Questionnaire and electronic consent data are entered by the participants and will include participant identifiable information to minimise the burden on research staff, as the sample size has increased. This will be managed through the SIREN information assets, with data received through SnapSurvey (this is a secure web-based system hosted on the PHE server) and imported into the SIREN SQL server.
- In addition to the participant questionnaires, vaccination data may be collected via the National Immunology Management System (NIMS). This includes date of vaccination, manufacturer and batch number of the vaccine.
- Public Health England and the Devolved Administration Public Health Agencies may access centrally held information to find identifiers (e.g. NHS number) and may also link centrally held health and care data to the data collected in SIREN.
- Viral genomic data generated by the local or consortium laboratories, identified by the Consortium identifier (pseudonymised with the link to personal information held by PHE under existing surveillance protocol). Genomic data for COG-UK is held within PHE data assets and within the PHE network.
- Viral culture results and the results of PHE serology testing will be available to the PHE SIREN team through linkage with the PHE MOLIS data asset.

All source data will be securely transferred to PHE with the identifiers described above. Data will be received into PHE from the survey (a secure and encrypted web survey hosted on a PHE server), via secure email (recruitment log), through PHE's established data assets Second Generation Surveillance System (SGSS; laboratory data), MOLIS, or through end to end encrypted bespoke arrangements (e.g. genomic data). The person identifiable information is required to run the follow-up survey (dependent on having names and contact number/email), and for PHE to undertake secure data linkage across the data assets as well as drawing in data from the Lighthouse laboratories.

The SIREN database will be a SQL database (with MS Access front-end) on a secure server at PHE with access restricted to named authorised staff within the PHE SIREN team. Participant Identifiable data including full name, email address and mobile number is stored in a separate database table to the main parent table. This will reduce viewing frequency of this data as the main parent table will be regularly accessed by the study team to generate reports. All analysis will be performed by study personnel on the pseudonymised linked datasets.

**Access to Data:** Direct access will be granted to authorised representatives from the Sponsor for monitoring and/or audit of the study to ensure compliance with regulations. Access to study data and responsibility for undertaking data management and analysis will be undertaken by named authorised staff within the PHE SIREN team. All members of the SIREN team involved in data linkage have been trained in handling data according to Caldicott guidelines and Section 60 of the Health and Social Care Act. All researchers are aware of the Data Protection Act 2018, and the need to maintain absolute confidentiality.

**Data Recording and Record Keeping:** The data will be securely held at the National Infection Service, PHE. Data collection, storage and use will be consistent with the procedures described in the NHS Information Governance Toolkit. All databases will be encrypted and appropriately access-restricted.

**Data storage:** Electronic data will be stored on PHE secure servers and will remain active for the duration of the study. Participants' identifiable data will then be removed and the data will be archived within the platform to be retained for a period of 5 years.

**Records retention policy:** The continued secure storage of participant data in all SIREN information assets will be subject to a regular (quarterly) review by the SIREN Study Management Group. Decisions on whether to continue to retain this data and the justification for this will be documented in the SIREN study Decision Log.

# 5. Analysis plan

## 5.1 Overall Analysis

All enrolled participants will be included in analyses, which will account for clustering by research site. Analyses will be conducted after each 4 week period to inform the UK's response to the COVID-19 pandemic. Results will be available to all organisations involved in the research. The

study follow-up period will end by default 12 months following the enrolment of the last participant, but by consensus of the study management group and funder may be terminated sooner if findings are sufficient.

There are no formal stopping rules for futility, utility or lack of power. The final decision to terminate the study will be made by Public Health England and Department for Health and Social Care.

Estimates of both cumulative incidence and incidence density in the seropositive and seronegative cohorts will be obtained using mixed effects models assuming counts of PCR positive have a negative binomial distribution, a log link function, and the natural logarithm of the total number of subjects or the total follow-up time use as an offset, respectively. Inclusion of a binary predictor indicating the sero-status of the cohort into this model will provide estimates of the incidence rate ratio. Sites will be incorporated as a random intercept to account for unmeasured, shared, site level factors. To account for a non-constant force of infection, calendar month will be incorporated as an additional random effect. An assessment of the role of factors such as age, gender, ethnicity in immunity will be explored by inclusion of interactions within the model between each and serological status.

While the above analytical approaches provide a "classical" person-years approach to prospective cohort analysis and provide familiar measures of association, it may be inadequate to assessment of immunity provided by seroconversion. As it is expected that seropositivity is likely to confer a degree of short to median term protection for a SARS-CoV-2 infection, multi-state and parametric cure rate models incorporating frailty will be employed. These "survival" type of models provide a more detailed assessment of factors associated with both short term and longer-term protection from infection, and how immunity may wane over time. Both mixture, explicitly assuming an immune and non-immune group and non-mixture "cure rate" models will be assessed using information criterion to choose which provides a better fit to the observed data. Bayesian approaches to cure rate models with frailty as describe by deSouza[13] will be employed.

Multi state models explicitly allowing those within an "immune" state to flow into a "susceptible" state as antibodies wane will also be employed. This framework can allow subjects to move from seronegative (susceptible) to seropositive (immune) when infected during the study period. An additional absorbing state will be used for those infected that died. It is also possible to introduce "misclassification" of state into the multi state model, providing an estimate of sensitivity to account for imperfect serological tests. Approaches like those proposed by Jackson[14] will be employed.

5.2 Analysis for vaccine effectiveness

Survival analysis will be used to estimate the hazard ratio in vaccinated compared to unvaccinated SIREN participants with VE = 1 – HR. A nested test negative case-control analysis will also be done with those swabbed but negative as the controls. If more than one vaccine is used vaccine effectiveness will be stratified by vaccine manufacturer. Vaccine effectiveness will also be stratified by baseline positivity (either PCR or antibody), age group (<50, >=50) and time since vaccination (three-month intervals and as a spline). Interaction with sex, ethnicity and risk group will be tested and, if significant, vaccine effectiveness will be stratified by these factors.

If the vaccine is rolled out over a very short period to HCWs with very high coverage then the unvaccinated group will be small and probably an unusual subset. Even if coverage is not high those that do not get vaccinated when it is highly recommended may be different in ways that could lead to confounding. For example, those previously infected may not see the need for vaccination, or those not regularly working on site might miss vaccination. Those that perceive themselves as low risk of severe disease or with less patient contact may also be less likely to get the vaccine. Those not getting vaccinated may also be more likely to be those not providing regular swabs or blood samples. It will therefore be important to compare the vaccinated and unvaccinated cohorts to identify these potential biases. Using only those completing regular follow-up may help reduce such biases.

If coverage is very high and rapid then instead of vaccine effectiveness assessment it may be possible to do an impact assessment using a controlled interrupted time series approach in which COVID-19 incidence is compared over time in the HCW population to the general population (using external data) or between sites if vaccine introduction varies sufficiently by site. This can be done using Poisson or negative binomial regression.

**Procedure for Accounting for Missing, Unused, and Spurious Data:** Analyses will be restricted to cases with antibody and PCR tests. The PCR test for virus is being used as a diagnostic test and hence has high performance. Sufficient sera will be obtained to re-run the immunological assays in case of initial assay failure. For similar reasons we do not anticipate that spurious data will be obtained.

**Procedures for Reporting any Deviation(s) from the Original Statistical Plan:** Deviations from the original statistical plan or the statistical analysis plan will be described and justified in the analysis reports.

# 6. Confidentiality and information governance

The data management plan (Section 4) demonstrates that PHE and the Public Health bodies in the Devolved Administrations will need to be able to extract data from multiple healthcare organisations, including staff moving between organisations, national testing laboratories undertaking clinical diagnostic tests when staff are unwell, and potentially national sequencing consortium laboratories. Data linkage based on a single study identifier is neither practical nor robust in such circumstances. The suggested process for handling identifying data in a secure and appropriate way is therefore as follows:

- On enrolment, participants will be given a unique study number (three letter/number combination (five for Scottish Health Boards and Northern Irish Health and Social Care Trusts) identifying the healthcare organisation, plus five digits identifying the participant)
- The healthcare organisation will compile a participant database (Recruitment Log) containing the name, job role, date of birth, sex NHS number (or equivalent unique

identifier in Devolved Administrations), and study number, passcode of each enrolled participant. They may need to collect other personal information in order to monitor quotas, request tests and contact participants appropriately. They may also develop local tools to assist managing patient, sample and data flows and returning results; if they do so this should be stored according to information governance requirements.

- The local healthcare organisation will retain these data for local study management and will transfer a copy of the database securely to PHE using appropriate encryption if requested and will be stored, password protected, in a limited-access environment on a PHE server.
- PHE and the Public Health Organisations in the Devolved Administrations will use the participant's personal information to retrieve their test results from the national surveillance systems as described in the data management plan. They may also use personal information to request the residual samples for additional processing in the sub studies and to link to identifiers and health and care data held centrally.
- Positive swabs from the study will be routinely sequenced under the arrangement between the NHS and the COG-UK Consortium. In order to retrieve the genomic data, PHE and Public Health Organisations in the Devolved Administrations may use the participant's identifying information to search the COG-UK data hub, which is **within** the public health organisations, and retrieve the COG-UK study ID. The COG-UK study ID allows the public health organisations to retrieve the viral genomes from that participant's sample from COG-UK (which maintains a pseudonymised genome collection on a non-PHE server). The SIREN study team will make a request to the COG-UK team and will not be able to access any other fields in the COG-UK data hub or access the data hub directly.
- The SIREN SQL database will receive and manage data from the above data sources. This will include participant identifiable information, as required for testing linkage and communication of follow-up survey links to the participants. The continued retention of participant identifying information will be subject to regular review by the Study Management Group, as per the Records Retention policy.
- If a participant moves to another healthcare organisation, it will be the responsibility of the original enrolling organisation to organise the participant’s transfer to a new SIREN participating organisation, and to notify the PHE SIREN team of this change. Any communication involving PII or study number should be sent by secure email, and guidance is provided to all healthcare organisations on this.
- If, following review of results, the study team deem it necessary for result interpretation, access to past results or clinical management of a participant, we may contact sites to discuss individual participant results either by email or on the telephone; the discussion would be with the most appropriate individual at the site e.g. research team, occupational health, microbiology consultant. We may request access to pseudonymised results which are usually held by Occupational Health.

PHE has undertaken some seroprevalence surveys (LondonCOVID and ESCAPE); these studies have been/are being terminated and individuals who are eligible are being offered the opportunity to transition into SIREN. As the data from these studies would be extremely useful to SIREN, we are keen to access these results, however the protocol/consent for these surveys do not include the use of patient identifiable information by PHE or the use of results in other studies. Therefore, we have included a supplementary consent form for these participants which the local research

teams will retain locally. The use of data from these surveys is optional and refusing consent would not impact and individual from being able to participate in SIREN.

# 7. Quality assurance and research governance

The study may be monitored or audited in accordance with the study protocol and standard operating procedures, GCP and relevant regulations.

**Risk assessment:** No formal risk assessment is required. The study involves recruiting individuals without symptoms who will be asked to give full informed consent to have swabs taken or self-swab their nose or nose and throat, and provide a blood sample taken by a study research nurse, the hospital phlebotomy department or another individual judged to have been appropriately trained. Further participation in follow up visits to collect the same samples is based on consent of the participant. The main burden of participating in the study is the time taken for the study visits (which will be as far as possible at the same time as routine organisation mandated screening if it is in place) or the potential for minimal bruising from blood sampling though this is unlikely with experienced staff taking the samples. As to the potential impact on work if a sample is returned as positive; while a positive result would necessitate self-isolation, the impact of identifying these asymptomatically positive cases would be regarded as a positive outcome from the perspective of reduced of risk to patients, other staff and social contacts. There is minimal risk of harm to any patient from participating since it does not include any therapeutic intervention. The diagnostic test for the presence of virus from the nose or nose and throat swab will be conducted by an accredited laboratory (or those with satisfactory verification documents) and will be returned to the participant's research nurse and participant through standard hospital practice.

**Study monitoring:** No GCP monitoring will be undertaken. As described there are minimal risks posed to patients by this observational and non-interventional study. The only study procedures are the completion of the questionnaire and taking of samples.

**Safety reporting:** There are no interventions in this study, and the only procedures are a standard blood draw performed by a healthcare professional, and a participant swab by a trained professional or self-swab using a methodology that is being used widely across the country. Therefore, there is minimal safety risk to participants and safety reporting is not applicable.

**Study committees:** Oversight will be provided by a study management group including the investigators named above, representatives of collaborating and participating organisations as appropriate, and chaired by the Chief Investigator.

**Protocol deviations**: A deviation is a departure from the approved study protocol or other study document or process, or from Good Clinical Practice or any applicable regulatory requirement. Any deviations from protocol will be documented by the team delivering SIREN in each organisation in a protocol deviation spreadsheet, which will be stored securely and sent to PHE securely on a monthly basis.
Protocol deviations which we would expect to be recorded and reported to the Sponsor monthly include (but are not limited to):

- Participants who miss more than eight consecutive visits.
- Participants who did not have the correct samples undertaken at any given visit
- Where samples are incorrectly discarded or not tested as per protocol

**Serious breaches**: A serious breach is a breach of the protocol or of the conditions or principles of Good Clinical Practice which is likely to affect to a significant degree –

(a) the safety or physical or mental integrity of the trial subjects; or

(b) the scientific value of the research.

In the event that a serious breach is suspected the Sponsor must be contacted within 1 working day. In collaboration with the Chief Investigator, the serious breach will be reviewed by the Sponsor and, if appropriate, the Sponsor will report it to the approving REC committee and the relevant local healthcare organisation within seven calendar days. A serious breach would include a data breach. The organisation (both local and relevant public health body) will be expected to complete a review document which examines and circumstances, response, root cause and mitigation actions, and return this document to the Sponsor.

# Appendix 1: Wider Immunological Studies

| Study Name | Chief Investigator(s) | Study Aims |
|---|---|---|
| **PITCH**<br><br>**P**rotective **I**mmunity from **T** cells to **C**ovid-19 in **H**ealth workers | Paul Klenerman<br>Susanna Dunachie | The overall aim of this study is to determine the relationship between T cell immunity and new SARS-COV-2 infection in seronegative and seropositive healthcare workers (HCW).<br><br>Full Protocol and Consent form (SIREN _PITCH) |
| **HICC**<br><br>**H**uman **I**mmune **C**orrelates for **C**OVID-19 | Helen Baxendale<br>Wilhelm Schwaeble<br>Jonathan Heeney | HICC is funded by UKRI and NIHR to support SIREN and other UK-wide COVID-19 studies to determine the humoral correlates of immunity against SARS-CoV-2.<br><br>The collaboration between SIREN and HICC aims to explore in detail the specificity, characteristics, and functionality of the antibody-based response to SARS-CoV-2 infection, re-infection, and vaccination within healthcare workers, with a special focus on memory B-cell responses.<br><br>A subset of SIREN samples will be shared with the HICC team. Testing is covered by the existing SIREN consent process.<br><br>SIREN-HICC lab protocol |
| **Oxford**<br><br>**University**<br><br>**Hospitals** | David Eyre<br>Katie Jeffery | Surveillance |

# Appendix 2: Amendment History

| Amendment No. | Protocol Version No. | Date issued | Author(s) of changes | Details of Changes made |
|---|---|---|---|---|
| 1 | 2 | 20052020 | Susan Hopkins | Removed Appendix 1 including PIL, Consent and Questionnaires<br><br>Added table of contents |
| 2 | 3.1 | 08062020 | Susan Hopkins | Modified Consent to be electronic to reduce data burden.<br><br>Increased sample size to determine whether there is immunity at much shorter time intervals to provide UK government with evidence of immunity or not at much faster intervals.<br><br>Modified sample storage to reflect the larger sample size<br><br>Added that PCR swab samples could form part of screening protocol for Trust or specific to this study and taken by nose and throat swab or nose self sampling.<br><br>Removed RT-PCR and changed to PCR to allow other local platforms to be used. |

| | | | | |
|---|---|---|---|---|
| **3** | **4.1** | **16082020** | **Susan Hopkins** | **Eligibility for sites extended to primary care/CCG and the independent sector.**<br><br>**Eligibility for sites from Devolved Administrations added**<br><br>**Eligibility of PIC (Participant identification centres) sites**<br><br>**Further details on inclusion/exclusion criteria.**<br><br>**Further details on electronic consent.**<br><br>**Action on positive results expanded**<br><br>**Removal of mention of NHS policy for routine swab testing in all hospitals in England.**<br><br>**Removal of plan to provide quotas initially.**<br><br>**Clarification of recruitment period to end 30/9/2020 for all sites.**<br><br>**The enrolment log to continue to be maintained by local organisation but transferred to PHE only on request.**<br><br>**Change from multiple aliquots of follow-up serum to one full 2ml serum.**<br><br>**Clarification of which follow-up serum samples need stored and sent to PHE.**<br><br>**Viral culture detail added.**<br><br>**Option of requirement to obtain additional swab for further testing.**<br><br>**Further description of who can take blood/swab samples – can be done by peer to peer sampling as long as** |

| | | | | |
|---|---|---|---|---|
| | | | | **approved by organisation guidance and individuals appropriately trained.**<br><br>**Added using individual's results from previous studies (including additional consent procedures for those in sero-prevalence surveys).**<br><br>**Options for publicity materials added**<br><br>**Description for paper consent process if website is down.**<br><br>**Further description of staff who are eligible**<br><br>**Protocol Deviations further defined with reporting mechanisms for deviations and breaches**<br><br>**Explicitly stated that local teams will need access to participants electronic record.**<br><br>**Further information about data collection and storage**<br><br>**Records retention policy added** |

| | | | | |
|---|---|---|---|---|
| **4** | **5.0** | **20112020** | **Susan Hopkins** | **Date of recruitment extended to 31/03/21.**<br><br>**SARS-COV2 updated to SARS-CoV-2**<br><br>**Reference to wider immunology studies added**<br><br>**Appendix of wider immunology collaboration studies added**<br><br>**Maximum volume of blood samples increased to facilitate sample collection for wider immunology studies**<br><br>**Data Protection Act 1998 updated to 2018 to include GDPR legislation**<br><br>**Updated protocol deviations to reduce burden on sites** |
| **5** | **5.1** | **23122020** | **Susan Hopkins** | **Updates to reflect the introduction of a COVID-19 vaccine among healthcare workers, including; updated background, two aims added, updated analysis plan and the expansion of study data to include data from the National Immunology Management System (NIMS).**<br><br>**Updated additional serology testing and assays to include analysis of B cells.**<br><br>**Additional collaboration study included in Appendix 1.** |
| **6** | **5.2** | **07012021** | **Susan Hopkins** | **Leaflet at vaccination added as an additional means of recruitment** |
| **7** | **6.0** | **28052021** | **Susan Hopkins** | **Updates to reflect the addition of an extended follow up period including; updated summary, background, primary aim, enrolment and consent, data collection and methods and withdrawal procedures.** |

| | | | | |
|---|---|---|---|---|
| **8** | **6.1** | **17062021** | **Susan Hopkins** | **Details of optional antibody testing added as an additional testing method, subject to local procedures.**<br><br>**Appendix name amended to 'Wider Immunological Studies' for clarity**<br><br>**Frequency of missed visits increased from three to eight missed visits to qualify as a protocol deviation.** |
| **10** | **6.2** | **11/08/2021** | **Susan Hopkins** | **Addition of paragraph outlining retention strategy and enabling participants to receive tokens of appreciation** |
| | | | | |

# References


1. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020;382(8):727-33.
2. World Health Organisation. WHO announces COVID-19 outbreak a pandemic. Available from: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic. WHO: Geneva, 2020 (accessed 11 May 2020).
3. World Health Organisation. Coronavirus disease (COVID-19) Situation Dashboard. Available from: https://covid19.who.int/. WHO: Geneva, 2020 (accessed 11 May 2020).
4. Public Health England. Guidance and standard operating procedure COVID-19 virus testing in NHS laboratories. Available from: https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/guidance-and-sop-covid-19-virus-testing-in-nhs-laboratories-v1.pdf. PHE: London, 2020 (accessed 11 May 2020).
5. Konrad R, Eberle U, Dangel A, et al. Rapid establishment of laboratory diagnostics for the novel coronavirus SARS-CoV-2 in Bavaria, Germany, February 2020. *Euro Surveill* 2020;25:9.
6. To KK, Tsang OT, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020;20(5):565-574.
7. Wikramaratna P, Paton RS, Ghafari M, Lourenco J. Estimating false-negative detection rate of SARS-CoV-2 by RT-PCR. *MedRvix* 2020; https://www.medrxiv.org/content/10.1101/2020.04.05.20053355v2**.**
8. Li Z, Yi Y, Luo X, et al. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. *J Med Virol* 2020;doi:10.1002/jmv.25727
9. Hamre D, Beem M: Virologic studies of acute respiratory disease in young adults. V. Coronavirus 229E infections during six years of surveillance. *Am J Epidemiol.* 1972; 96: 94–106.
10. Hamre D, Beem M: Virologic studies of acute respiratory disease in young adults. V. Coronavirus 229E infections during six years of surveillance. *Am J Epidemiol.* 1972; **96**: 94–106.
11. Treibel TA, Manisty C, Burton M. et al. COVID-19: PCR screening of asymptomatic healthcare workers at London hospital. *Lancet* 2020;18(1):462.

12. Department of Health and Social Care. Priority groups for coronavirus (COVID-19) vaccination: advice from the JCVI, 2 December 2020. Available from https://www.gov.uk/government/publications/priority-groups-for-coronavirus-covid-19-vaccination-advice-from-the-jcvi-2-december-2020/priority-groups-for-coronavirus-covid-19-vaccination-advice-from-the-jcvi-2-december-2020 DHSC: London, 2020 (accessed 7 December 2020).
13. de Souza D, Cancho VG, Rodrigues J, Balakrishnan N. Bayesian cure rate models induced by frailty in survival analysis. *Stat Methods Med Res* 2017;26(5):2011-28
14. Christopher H. Jackson. Multi-State Models for Panel Data: The msm Package for R. *J Stat Soft* 2011;38:8.

SIREN

SARS-CoV2 Immunity & Reinfection Evaluation

# SIREN Study Statistical Analysis Plan

| Sections | Version | Date |
|---|---|---|
| **1. Trial Statistical Analysis plan**<br>• As recorded in original study protocol | 1.0 | June 2020 |
| **2. Current Statistical Analysis plan**<br>• Vaccine Effectiveness Statistical Plan<br>• First interim analysis<br>• First vaccine effectiveness analysis<br>• Second interim analysis | 3.3 | September 2021 |
| **3. Summary of changes to Statistical Plan** | N/A | September 2021 |

Signed Study Statistician 

Dr Andre Charlett

Date 1 September 2021

Signed Chief Investigator 

Dr Susan Hopkins

Date 1 September 2021

## 1. Trial statistical analysis plan, June 2020

All enrolled participants will be included in analyses, which will account for clustering by research site. Analyses will be conducted after each 4-week period to inform the UK's response to the COVID-19 pandemic. Results will be available to all organisations involved in the research. The study follow-up period will end by default 12 months (unless additional funding achieved) following the enrolment of the last participant, but by consensus of the study management group and funder may be terminated sooner if findings are sufficient.
There are no formal stopping rules for futility, utility or lack of power. The final decision to terminate the study will be made by Public Health England and Department for Health and Social Care.
Estimates of both cumulative incidence and incidence density in the seropositive and seronegative cohorts will be obtained using mixed effects models assuming counts of PCR positive have a negative binomial distribution, a log link function, and the natural logarithm of the total number of subjects or the total follow-up time use as an offset, respectively. Inclusion of a binary predictor indicating the sero-status of the cohort into this model will provide estimates of the incidence rate ratio. Sites will be incorporated as a random intercept to account for unmeasured, shared, site level factors. To account for a non-constant force of infection, calendar month will be incorporated as an additional random effect. An assessment of the role of factors such as age, gender, ethnicity in immunity will be explored by inclusion of interactions within the model between each and serological status.
While the above analytical approaches provide a “classical” person-years approach to prospective cohort analysis and provide familiar measures of association, it may be inadequate to assessment of immunity provided by seroconversion. As it is expected that seropositivity is likely to confer a degree of short to median term protection for a SARS-CoV-2 infection, multi-state and parametric cure rate models incorporating frailty will be employed. These “survival” type of models provide a more detailed assessment of factors associated with both short term and longer-term protection from infection, and how immunity may wane over time. Both mixture, explicitly assuming an immune and non-immune group and non-mixture “cure rate” models will be assessed using information criterion to choose which provides a better fit to the observed data. Bayesian approaches to cure rate models with frailty as describe by deSouza[1] will be employed.
Multi state models explicitly allowing those within an “immune” state to flow into a “susceptible” state as antibodies wane will also be employed. This framework can allow subjects to move from seronegative (susceptible) to seropositive (immune) when infected during the study period. An additional absorbing state will be used for those infected that died. It is also possible to introduce “misclassification” of state into the multi state model, providing an estimate of sensitivity to account for imperfect serological tests. Approaches like those proposed by Jackson[2] will be employed.

**Procedure for Accounting for Missing, Unused, and Spurious Data:** Analyses will be restricted to cases with antibody and PCR tests. The PCR test for virus is being used as a diagnostic test and hence has high performance. Sufficient sera will be obtained to re-run the immunological assays in case of initial assay failure. For similar reasons we do not anticipate that spurious data will be obtained.

**Procedures for Reporting any Deviation(s) from the Original Statistical Plan:** Deviations from the statistical analysis plan will be described and justified in the analysis reports.

### References

[1] Souza, Daiane & Cancho, Vicente & Rodrigues, Josemar & Balakrishnan, Narayanaswamy. (2017). Bayesian cure rate models induced by frailty in survival analysis.

Statistical Methods in Medical Research. 26. 096228021770867. 10.1177/0962280217708671.

[2] Jackson, C. (2011). Multi-State Models for Panel Data: The msm Package for R. *Journal of Statistical Software*, *38*(8), 1–28. https://doi.org/10.18637/jss.v038.i08

## 2. Current SIREN Statistical Analysis Plan Version 3.3

All enrolled participants will be included in analyses, which will account for clustering by research site. Analyses will be conducted after each 4-week period to inform the UK's response to the COVID-19 pandemic. Results will be available to all organisations involved in the research. The study follow-up period will end by default 12 months (unless additional funding achieved) following the enrolment of the last participant, but by consensus of the study management group and funder may be terminated sooner if findings are sufficient.

There are no formal stopping rules for futility, utility or lack of power. The final decision to terminate the study will be made by Public Health England and Department for Health and Social Care.

Estimates of both cumulative incidence and incidence density in the seropositive and seronegative cohorts will be obtained using mixed effects models assuming counts of PCR positive have a negative binomial distribution, a log link function, and the natural logarithm of the total number of subjects or the total follow-up time use as an offset, respectively. Inclusion of a binary predictor indicating the sero-status of the cohort into this model will provide estimates of the incidence rate ratio. Sites will be incorporated as a random intercept to account for unmeasured, shared, site level factors. To account for a non-constant force of infection, calendar month will be incorporated as an additional random effect. An assessment of the role of factors such as age, gender, ethnicity in immunity will be explored by inclusion of interactions within the model between each and serological status.

While the above analytical approaches provide a “classical” person-years approach to prospective cohort analysis and provide familiar measures of association, it may be inadequate to assessment of immunity provided by seroconversion. As it is expected that seropositivity is likely to confer a degree of short to median term protection for a SARS-CoV-2 infection, multi-state and parametric cure rate models incorporating frailty will be employed. These “survival” type of models provide a more detailed assessment of factors associated with both short term and longer-term protection from infection, and how immunity may wane over time. Both mixture, explicitly assuming an immune and non-immune group and non-mixture “cure rate” models will be assessed using information criterion to choose which provides a better fit to the observed data. Bayesian approaches to cure rate models with frailty as describe by deSouza[1] will be employed.

Multi state models explicitly allowing those within an “immune” state to flow into a “susceptible” state as antibodies wane will also be employed. This framework can allow subjects to move from seronegative (susceptible) to seropositive (immune) when infected during the study period. An additional absorbing state will be used for those infected that died. It is also possible to introduce “misclassification” of state into the multi state model, providing an estimate of sensitivity to account for imperfect serological tests. Approaches like those proposed by Jackson[2] will be employed.

**Procedure for Accounting for Missing, Unused, and Spurious Data:** Analyses will be restricted to cases with antibody and PCR tests. The PCR test for virus is being used as a diagnostic test and hence has high performance. Sufficient sera will be obtained to re-run the immunological assays in case of initial assay failure. For similar reasons we do not anticipate that spurious data will be obtained.

**Procedures for Reporting any Deviation(s) from the Original Statistical Plan:** Deviations from the statistical analysis plan will be described and justified in the analysis reports.

**Vaccine Effectiveness analysis plan December 2020**

Survival analysis will be used to estimate the hazard ratio in vaccinated compared to unvaccinated SIREN participants with VE = 1 – HR. A nested test negative case-control analysis will also be done with those swabbed but negative as the controls. If more than one vaccine is used vaccine effectiveness will be stratified by vaccine manufacturer. Vaccine effectiveness will also be stratified by baseline positivity (either PCR or antibody), age group (<50, >=50) and time since vaccination (three-month intervals and as a spline). Interaction with sex, ethnicity and risk group will be tested and, if significant, vaccine effectiveness will be stratified by these factors.

If the vaccine is rolled out over a very short period to HCWs with very high coverage then the unvaccinated group will be small and probably an unusual subset. Even if coverage is not high those that do not get vaccinated when it is highly recommended may be different in ways that could lead to confounding. For example, those previously infected may not see the need for vaccination, or those not regularly working on site might miss vaccination. Those that perceive themselves as low risk of severe disease or with less patient contact may also be less likely to get the vaccine. Those not getting vaccinated may also be more likely to be those not providing regular swabs or blood samples. It will therefore be important to compare the vaccinated and unvaccinated cohorts to identify these potential biases. Using only those completing regular follow-up may help reduce such biases.

If coverage is very high and rapid then instead of vaccine effectiveness assessment it may be possible to do an impact assessment using a controlled interrupted time series approach in which COVID-19 incidence is compared over time in the HCW population to the general population (using external data) or between sites if vaccine introduction varies sufficiently by site. This can be done using Poisson or negative binomial regression.

**Key subsequent decisions made concerning statistical analysis methods**

First interim analysis

Participants will be assigned to the positive cohort if they meet one of the following criteria: antibody positive on enrolment or antibody positive from prior clinical laboratory sample, with or without prior PCR positive; antibody negative on enrolment with prior antibody positive, with or without prior PCR positive; antibody negative on enrolment with a PCR positive result prior to enrolment.

Participants will be assigned to the negative cohort if they had a negative antibody test and no documented positive PCR test. Those in the negative cohort move to the positive cohort 21 days following a PCR positive test result or at the time of antibody seroconversion with no positive PCR test.

The SIREN case definitions for reinfections will be used.

The cohort will be described by their baseline cohort allocation. Cumulative incidence, using the total number of participants in each cohort, and incidence density using the total person time at risk will be calculated for both cohorts and sub-categories and plotted over time using PCR confirmation only. A mixed effects logistic regression analysis will be used to estimate odd ratios to measure the association between the exposure (cohort allocation) and the binary outcome (PCR test result). The entry date used in this analysis for all participants will be the earliest antibody test. All PCR tests after the entry date will be used, though the reinfection case definitions will account for multiple PCR test positivity post primary infection. To account for temporal changes in the background risk of infection, tests will be allocated to the calendar week of the test date. Study site will be fitted as a random effect to account for the longitudinal nature of the study data, with age group, gender, ethnicity, staff

group, and region fitted as non-time varying fixed effects to account for their possible confounding effect. Analysis will be conducted in STATA v15.1 (College Station, TX: StataCorp LLC).

**First Vaccine Effectiveness Analysis**

Participants will be excluded from this analysis if they enrolled after 7 December 2020, had no PCR tests after 7 December 2020, or had insufficient PCR and antibody data to complete cohort assignment.
The primary outcome variable for the vaccine coverage analysis will be the binary 'ever vaccinated' variable. Participants will be categorised as 'ever vaccinated' if they had at least one vaccine dose recorded from 8 December 2020 to 5 February 2021 from at least one of the two vaccination data sources available.
The primary outcome variable for the vaccine effectiveness analysis will be a PCR confirmed SARS-CoV-2 infection. This is defined as a new PCR positive result during follow-up for the negative cohort and a reinfection during the follow-up in the positive cohort, irrespective of symptom status. Participants will be assigned into either the positive cohort (antibody positive or history of infection (prior antibody or PCR positive)) or the negative cohort (antibody negative with no prior positive test) at the beginning of the follow up period (7 December 2020). Follow-up time for all participants will start on 7 December 2020, the day before vaccine roll-out began, with all participants contributing at least one day of follow-up unvaccinated. Participants will move from unvaccinated to vaccinated within their assigned cohort on the date of the first vaccination dose. Participants will contribute person-time to follow-up until either an event of interest (i.e. a new PCR positive in the negative cohort or a reinfection in the positive cohort); the date of the suspect second dose for those with an unreliable date of second dose; the date of their first dose for those vaccinated with the ChAdOx1 vaccine; or the censored date. We will define the end of follow-up in those who were not positive cases as the date of a negative test or 05 February 2021 if the test was after this date, in order to avoid immortal time bias.
Investigation of factors associated with vaccination will be conducted using mixed effect multivariable logistic regression model (with hospital site as a random effect) to investigate confounding between demographic and occupational risk factors on the outcome variable 'ever vaccinated'. Only the variables which demonstrated strong evidence of association on vaccine coverage will be retained in the reported model.
A mixed effect proportional hazards frailty model using a Poisson distribution will be used to estimate Hazard Ratios to compare time to infection in unvaccinated and vaccinated participants to estimate the impact of the BNT162b2 vaccine on infection (including asymptomatic and symptomatic as the primary outcome). Hazard ratios from 21 days after first dose and seven days after second dose will be calculated using a weighted average method, the point at which an immunological response to the vaccine dose should have been provoked. Vaccine effectiveness was calculated as 1 – adjusted Hazard Ratio (vaccinated versus unvaccinated).
Three models will be run on different cohorts within the study population. The main model will include the full study population and adjusted for cohort assignment. Models will also be run on the two cohorts separately, to provide estimates of vaccine effectiveness in the susceptible population (negative cohort) and the positive cohort with natural immunity following prior SARS-CoV-2 infection.

**Second interim analysis**

Participants will be assigned into one of two cohorts at the start of analysis time: participants in the naïve cohort had no history of SARS-CoV-2 positivity and the positive cohort being those who had ever received a PCR positive or antibody result consistent with prior SARS-CoV-2 infection. Participants will be excluded from this analysis if event or cohort

assignment could not be accurately completed, i.e. no PCR tests during follow-up, or if they were in the positive cohort but were infected after vaccination or lacked an onset date for primary infection (PCR positive or COVID-symptom onset).
The primary outcome will be a PCR-confirmed SARS-CoV-2 infection, irrespective of symptom status, that met the definition of a reinfection in the positive cohort of two PCR positives ≥90 days or one new PCR positive ≥28 days after an antibody positive result consistent with previous infection.
Follow-up will begin on 07 December 2020, the day before COVID-19 vaccination was introduced to the UK, and continued until 21 September 2021, covering 10 calendar months. All participants enrolled on or before 07 December 2020 will contribute follow-up time from 07 December 2020 onwards. Participants who were enrolled after 07 December 2020 will contribute follow-up time from their enrolment date. Participants will move from the negative to positive cohort 90 days after a primary PCR positive date, if their primary infection was before vaccination, at which point they were considered at risk of reinfection.
We will use a Cox proportional hazards model with delayed entry, the outcome being time-to-infection with a positive PCR test. **We will use two main predictors** – vaccine status and previous infection status – both will be categorical and time-varying, grouping on the time to vaccination and divided into follow-up time into unvaccinated and 24 post-vaccination time intervals, with post-vaccination intervals categorised by manufacturer, dose and dosing interval, the latter to explore differences in protection in those receiving dose two closer in time to their first dose. We will also investigate the time since primary infection using four time-intervals: before primary infection (naïve), and then 3-9 months, 9-15 months and ≥15 months after primary infection. Vaccine effectiveness and protection from primary infection will be calculated as 1-HR, using robust variance estimates to guard against the potential for unmeasured confounders at trust level.
We will fit a main effects model, with no interactions between vaccine and primary infection status and an interaction model, not considering time since vaccination, brand and manufacturer, to focus on protection from primary infection over time by vaccine status. Both models will be fitted with and without additional time invariant covariates: age, ethnicity, co-morbidities, region, frequency of COVID-19 patient contact, patient-facing role, and workplace setting.
Independently, we will also an equivalent piecewise exponential proportional hazards model. We will use STATA software (version 15.1; StataCorp LLC, College Station, TX, USA) for all analyses. Results will be independently replicated in R (v. 4.1.1, survival package v.3.2-13). We will provide annotated code in a public facing github (https://github.com/SIREN-study/SARS-CoV-2-Immunity).

## References

[1] Souza, Daiane & Cancho, Vicente & Rodrigues, Josemar & Balakrishnan, Narayanaswamy. (2017). Bayesian cure rate models induced by frailty in survival analysis. Statistical Methods in Medical Research. 26. 096228021770867. 10.1177/0962280217708671.

[2] Jackson, C. (2011). Multi-State Models for Panel Data: The msm Package for R. *Journal of Statistical Software*, *38*(8), 1–28. https://doi.org/10.18637/jss.v038.i08

## 3. Summary of Changes to Statistical Plan

| Amendment No. | Statistical Plan Version No. | Date Issued | Author of changes | Summary of change |
|---|---|---|---|---|
| 1 | 2 | December 2020 | Andre Charlett and Susan Hopkins | Addition of analysis of vaccine effectiveness, reflecting the introduction of vaccines in healthcare workers. |
| 2 | 3.1 | December 2020 | Andre Charlett and Susan Hopkins | Details added for interim analysis, including selection criteria and details of a mixed effects logistic regression to determine association between exposure and PCR positivity. |
| 3 | 3.2 | February 2021 | Andre Charlett and Susan Hopkins | Details added outline the first analysis of vaccine effectiveness. This includes selection criteria and an overview of the three models used to examine factors associated with efficacy of COVID-19 vaccination. |
| 4 | 3.3 | September 2021 | Andre Charlett and Susan Hopkins | Overview of selection criteria and methods used for the second interim analysis added, examining vaccine status and previous infection history as a predictor of infection. |