







BMJ Open Impact of prior SARS-CoV-2 infection and COVID-19 vaccination on the subsequent incidence of COVID-19: a multicentre prospective cohort study among UK healthcare workers – the SIREN (Sarscov2 Immunity & REinfection EvaluationN) study protocol

Sarah Wallace ¹, Victoria Hall,¹ Andre Charlett ², Peter D Kirwan ^{1,3}, Michele Cole,¹ Natalie Gillson,¹ Ana Atti,¹ Jean Timeyin,¹ Sarah Foulkes,¹ Andrew Taylor-Kerr,¹ Nick Andrews ², Madhumita Shrotri,¹ Sakib Rokadiya,¹ Blanche Oguti,¹ Amoolya Vusirikala,¹ Jasmin Islam,¹ Maria Zambon,¹ Tim J G Brooks ¹, Mary Ramsay,¹ Colin S Brown,¹ Meera Chand,¹ Susan Hopkins ¹

To cite: Wallace S, Hall V, Charlett A, *et al*. Impact of prior SARS-CoV-2 infection and COVID-19 vaccination on the subsequent incidence of COVID-19: a multicentre prospective cohort study among UK healthcare workers – the SIREN (Sarscov2 Immunity & REinfection EvaluationN) study protocol. *BMJ Open* 2022;**12**:e054336. doi:10.1136/bmjopen-2021-054336

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-054336>).

SW and VH are joint first authors.

Received 08 July 2021
Accepted 04 May 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Victoria Hall;
Victoria.Hall@phe.gov.uk

ABSTRACT

Introduction Understanding the effectiveness and durability of protection against SARS-CoV-2 infection conferred by previous infection and COVID-19 is essential to inform ongoing management of the pandemic. This study aims to determine whether prior SARS-CoV-2 infection or COVID-19 vaccination in healthcare workers protects against future infection.

Methods and analysis This is a prospective cohort study design in staff members working in hospitals in the UK. At enrolment, participants are allocated into cohorts, positive or naïve, dependent on their prior SARS-CoV-2 infection status, as measured by standardised SARS-CoV-2 antibody testing on all baseline serum samples and previous SARS-CoV-2 test results. Participants undergo monthly antibody testing and fortnightly viral RNA testing during follow-up and based on these results may move between cohorts. Any results from testing undertaken for other reasons (eg, symptoms, contact tracing) or prior to study entry will also be captured. Individuals complete enrolment and fortnightly questionnaires on exposures, symptoms and vaccination. Follow-up is 12 months from study entry, with an option to extend follow-up to 24 months. The primary outcome of interest is infection with SARS-CoV-2 after previous SARS-CoV-2 infection or COVID-19 vaccination during the study period. Secondary outcomes include incidence and prevalence (both RNA and antibody) of SARS-CoV-2, viral genomics, viral culture, symptom history and antibody/neutralising antibody titres.

Ethics and dissemination The study was approved by the Berkshire Research Ethics Committee, Health Research Authority (IRAS ID 284460, REC reference 20/SC/0230) on 22 May 2020; the vaccine amendment was approved

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ As far as the authors are aware this is the largest longitudinal cohort study globally examining the question of reinfection with SARS-CoV-2 in a highly exposed population; the study should capture the vast majority of positive cases with regular asymptomatic PCR testing, widespread workplace Lateral Flow Testing and data infrastructure to capture all additional testing following symptoms and exposures.
- ⇒ Rich prospective data collection on reinfections and vaccine breakthroughs, including serial serum samples (pre-event and postevent), enabling analysis of correlates of protection.
- ⇒ The flexible design of this cohort study has enabled the study to adapt to examine vaccine effectiveness as vaccines were rolled out in the UK.
- ⇒ The study design allows detailed investigations of sub-groups within the main cohort, through external research collaborations.
- ⇒ Differences in demographics, general health and ongoing risk of exposure between healthcare workers and the general population mean that the results may not be fully generalisable to the UK population.

on 12 January 2021. Participants gave informed consent before taking part in the study. Regular reports to national and international expert advisory groups and peer-reviewed publications ensure timely dissemination of findings to inform decision making. **Trial registration number** ISRCTN11041050.

INTRODUCTION

SARS-CoV-2, a novel coronavirus which causes respiratory illness, was first identified in China in December 2019.¹ Following global spread of the virus, the WHO declared a national pandemic in March 2020.² Globally 315 million cases had been reported to the WHO by 13 January 2022, with 5 510 174 deaths attributed to COVID-19,³ and both the virus and the measures put in place to reduce the spread have led to significant economic and societal impacts. Whether and why individuals are re-infected with SARS-CoV-2 or infected after COVID-19 vaccination, the severity and transmissibility of these infections, how long infection-acquired and vaccine-acquired protection last and the impact of different SARS-CoV-2 variants remain crucial questions to inform the ongoing pandemic response.

The risk of reinfection for individuals who have previously had COVID-19 was poorly understood when this study commenced in June 2020. Early evidence of cases of reinfection was demonstrated initially in a handful of examples.^{4–10} Since then a number of papers, including from Sarscov2 Immunity & REinfection EvaluationN (SIREN), have demonstrated reduced risk of infection, generally over 80%, in those with prior infection.^{11–14} However, there is also evidence of reducing infection-acquired¹⁵ and vaccine-acquired immunity over time demonstrated by reinfections/breakthrough infections.^{15–17}

A large number of studies have examined antibody titres over time following both infection and vaccination; while neutralising antibodies do appear to persist for several months, many studies report waning of titres over time but often with large variations between individuals.^{17–19} However, much still needs to be understood regarding the implications of antibody titres on protection from becoming infected or suffering severe illness. Continued follow-up of well-defined cohorts is therefore enormously valuable, particularly given evidence of vaccine waning and with the emergence of new variants, most recently Omicron.²⁰

Surveys of healthcare workers in the UK have consistently demonstrated higher positive antibody prevalence compared with the general population,^{21–24} while one early observational cohort study of over 2 million people in the USA and UK in 2020 found a HR of 3.4 (95% CI 3.37 to 3.43) among healthcare workers reporting a positive test compared with non-healthcare workers, after adjustment to account for more frequent testing in healthcare workers.²⁵ Reasons for higher rates may include occupational exposure and increased testing through staff testing programmes and/or easier access to testing. Healthcare workers are therefore an ideal population for a study examining the risk of reinfection, as they have a high baseline antibody prevalence, relatively high ongoing SARS-CoV-2 exposure and undergo frequent routine asymptomatic SARS-CoV-2 screening.

The SIREN study, a large prospective cohort of healthcare workers undergoing regular testing, was initiated to examine the impact of SARS-CoV-2 antibodies on the risk

of infection with COVID-19. However, it has been and continues to be agile in response to the evolving questions of the pandemic, adapting to new developments, including responding to research needs connected to vaccine deployment and variant emergence.

METHODS AND ANALYSIS

Study design

This is a prospective longitudinal cohort study which has enrolled over 44 000 individuals between 18 June 2020 and 31 March 2021. Individuals are initially followed up for 12 months, however, an extended optional follow-up period of a further 12 months follow-up is offered to participants at sites participating in the extension with regular data collection. Enrolment took place between 18 June 2020 and 31 March 2021. All participants' follow-up will be completed by 31 March 2023.

Study objectives

The overall aim of this study is to determine if prior SARS-CoV-2 infection in healthcare workers confers future immunity to reinfection.

Primary objective

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 reinfection over short-term periods (reviewed monthly) and the next year.

Secondary objectives

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 1 January 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 comorbidities.
4. To measure the ability to culture viable virus from cases of reinfection diagnosed by PCR and whether those who 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 confirmed reinfection.
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 infection and symptomatic disease.
11. To monitor immune response to vaccination over time

Participants and recruitment

Population

The eligible population are staff members of healthcare organisations. Staff are recruited from healthcare organisations participating as SIREN sites, and all National Health Service (NHS) Trusts/Health Boards (organisations that manage hospitals) in England, Scotland, Wales and Northern Ireland have been invited to join.

Eligibility

A participant is eligible to join the study if they are a healthcare organisation staff member who works in a clinical setting where patients are present, can provide written consent, and is willing to remain engaged with follow-up for 12 months. Temporary short-term staff members are not eligible.

Recruitment and consent

Sites are responsible for recruiting eligible participants, according to their own processes. Sites are recommended to circulate all staff communications inviting volunteers and to monitor the demographics of their cohort as they recruit, aiming to represent their staff population. There are no requirements for quotas or structured sampling.

Interested and eligible potential participants are provided with a unique study number and passcode by their site research team and directed to enrol in the study by completing the online consent form and enrolment questionnaire. On completion of the online consent form and enrolment questionnaire, participants join the SIREN cohort. Site research teams are automatically informed of participant enrolment in real time and can then proceed with testing.

Recruitment into extended follow-up

Up to an additional 12-month follow-up is offered to participants at selected sites offering the study extension. Participants are sent an Extension Preference Survey 4 weeks before their original study end date, allowing them to opt in or out of continued follow-up. Site teams are automatically informed of participant responses to organise appropriate testing. For participants opting in, the Extension Preference Survey also collects updated demographic information, such as workplace setting and post code, in order to capture potential changes since completion of the enrolment survey at the beginning to their study period.

For sites opting out of the extension, participants do not receive the Extension Preference Survey and end their study period 12 months from enrolment.

Routine data collection and data management

At enrolment

At enrolment participants complete an online questionnaire, submit serum for SARS-CoV-2 antibody testing and a nose swab (or nose and throat swab, depending on local protocols) for SARS-CoV-2 antibody and nucleic acid amplification (NAAT) testing. Participants have up to 10mls of blood taken by venepuncture at enrolment and follow-up. The questionnaire collects information on participant demographics, work environment, symptom history, testing and vaccination history, participation in clinical trials and known COVID-19 exposures since 1 January 2020.

At follow-up

Participants undergo regular repeat NAAT and antibody testing throughout the study period, via nose/nose and throat swabs and serum samples respectively. NAAT testing is collected at fortnightly intervals and antibody testing at month intervals, although the protocol allows for testing frequency to be revised (weekly to monthly) subject to local/national epidemiology and feedback. For participants completing extended follow-up, from 12 to 24 months the default antibody testing frequency is quarterly but NAAT testing remains fortnightly.

Participants may also have additional PCR tests for other reasons which are outside the SIREN regimen, if for example they are symptomatic or identified as a contact of a case, and these test results are captured within SIREN (see data management).

Participants are sent a link to an online follow-up questionnaire on a fortnightly basis, with a reminder message sent after 2 days if the follow-up questionnaire is not completed. These questionnaires capture information on symptoms, exposures, vaccinations (both COVID-19 and seasonal influenza) and subsequent enrolment in vaccine or prophylaxis trials. For participants completing extended follow-up, they receive the same follow-up questionnaire at the same fortnightly frequency.

Testing at siren site laboratories

For all participants NAAT (typically PCR) and antibody testing is undertaken locally at the laboratory used by their healthcare organisation, this will be qualitative testing only (positive or negative). The healthcare organisation is responsible for issuing results to the participants as per local procedures. Testing platforms, including choice of antibody assay, is determined locally.

Data management

All laboratories for SIREN participating sites submit their antibody and antigen testing data into national laboratory surveillance systems. In England, this is the UK Health Security Agency's (UKHSA) Second Generation Surveillance System (SGSS), and there are equivalent

surveillance systems in the Devolved Administrations (Scotland, Wales and Northern Ireland). Testing data from sites on SIREN participants is obtained by the SIREN team through deterministic linkage, based on the NHS number (or equivalent unique identifier for Devolved Administrations) and other patient identifiers provided by participants in the enrolment questionnaire. Linkage with SGSS is undertaken daily and stored in the SIREN Structured Query Language database. Testing data from the Devolved Administrations is linked with the support of their respective public health agencies and the data are transferred via secure file transfer daily from the Devolved Administrations to the SIREN database. At enrolment, participants consent for the SIREN team to link all their historic and future SARS-CoV-2 testing data, including tests undertaken prior to enrolment, and tests taken outside SIREN, such as tests taken due to symptoms or exposures.

Vaccination status

Data on vaccination status, including dates vaccinated, dose, manufacturer and batch, is obtained directly from participants in the enrolment and follow-up questionnaires and through linkage on personal identifiers to the national COVID-19 vaccination registers. In England this is the National Immunisation Management System.

Enhanced data collection for investigation of events of interest

The UKHSA SIREN team run a daily query on the SIREN database, to identify any participants who are 'flagged' as an Event of Interest (EOI). EOIs include potential reinfections, defined as participants who had two positive PCR tests 90 days apart or antibody positive participants with a PCR positive test 4 weeks after their first antibody positive date, and vaccine breakthroughs (a new infection at least 21 days after first vaccine dose).

Once flagged, EOI cases are subject to a more detailed investigation, based on a survey which is sent to sites to obtain Ct values, confirm symptoms around the new infection episode and clarify sample location.

Since early spring 2021, all participants with a new PCR positive result are contacted to send them an additional postal self-swab in Viral Transport Media (VTM) to their home address, which is sent for PCR testing and sequencing at UKHSA Colindale laboratory.

Serum samples from EOI undergo enhanced serological testing at UKHSA Porton, as described below and results are captured in the SIREN Database.

Reinfections are classified as possible, probable, confirmed or excluded following review of the clinical, serological and genomic data.

Laboratory testing on siren samples

Serology

For all participants, at enrolment an aliquot of 2 mL serum will be shipped to and stored in the UKHSA SIREN biobank. At follow-up, serum samples for participants

who have ever been antibody positive or antigen positive, have received a COVID-19 vaccination or enrolled in a COVID-19 vaccine trial will be sent to and stored at the UKHSA SIREN biobank.

At enrolment, all participants will have their serum retested by UKHSA for antibodies to SARS-CoV-2, using the Roche Elecsys Anti-SARS-CoV-2 spike (S) and nucleocapsid (N) protein assays, providing quantitative assessment of antibody titres.²⁶ Individuals will be classified as seropositive or seronegative at baseline based on UKHSA antibody testing for N and S, with seropositivity to N used to identify those with previous SARS-CoV-2 infection.

Serological characterisation will be undertaken on additional participant samples for specific analyses. For example, cases of reinfection or vaccine failures, plus suitably matched controls, will have their sequential sera further characterised including anti-N and anti-S antibody titres (above) and additional assays including for the presence of neutralising antibody, to provide hypothesis generating data on mechanisms of protective immunity and correlates of immunity. Further detail on this additional testing aligned with specific analyses is provided elsewhere.¹⁰

To assess antibody trajectories over time subsets of participants of interest will be selected and their serial serum samples tested to conduct longitudinal cohort analyses.

Genomic analysis

Positive NAAT samples from participants from routine NAAT testing, should be sequenced as part of the routine sequencing of NHS SARS-CoV-2 positive samples. SARS-CoV-2 PCR positive EOI samples are sequenced in UKHSA Colindale or in an associated sequencing laboratory. Quality-controlled genomes are uploaded to the Cloud Infrastructure for Microbial Bioinformatics (CLIMB).

For participants who have more than one sequenced positive PCR sample potentially associated with separate infection episodes, 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 CLIMB, 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 reinfection or potential vaccine failure will be identified and viral culture requested for certain cases. This may be on residual VTM from the swab already taken at the trust, or on an additional swab in VTM.

T-Cell assays and other studies

Participants who have potentially been reinfected, are potential vaccine failures, are persistently NAAT positive, or have discordant serology may be contacted by the

SIREN Study Team to link into optional substudies, for example, assessing T cell assays and antibody dynamics.

Sample size and power

A simulation approach using a mixed effects Poisson regression model has been used to estimate the power to detect relative differences between the study cohorts. Our key assumptions include that 25% of our cohort will be seropositive at enrolment (based on 20% of staff who were asymptomatic and tested positive in one London hospital between 23 March and 2 May 2020,²³ and a total attrition of 35%, (unaffected by serostatus and occurring at a constant rate). The proportion of seropositive recruits at each site has been obtained from a Gaussian distribution with a mean of 0.25 and SD of 0.05 to reflect expected inter-site variation.

Power was estimated as the proportion of simulations for which the Wald statistic *p* value for the estimated incidence rate ratio in the seropositive compared with seronegative cohorts was less than 0.05. Our simulations found that there is statistical power of 80% or greater to detect a relative decrease of 30% or greater in cumulative incidence, provided the cumulative incidence in the seronegative group is in excess of 5%; even taking the cumulative incidence to as low as 2% in the seronegative group there is still sufficient power of in excess 80% for relative decrease of 80% or greater.

It was assumed that on average 250 participants would be recruited from each selected healthcare organisation, with a SD of 50. 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 represents a range 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 intertest period of 2 weeks.

It was assumed that the cumulative incidence in the seronegative cohort was 30% with a between trust coefficient of variation of 0.1, reflecting levels of seropositivity in HCWs at the time. Relative reductions in cumulative incidence in the seropositive cohort was varied between 1 (no protection from infection) to 0.1 (antibody effectiveness of 90%). Units in the simulations were 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 of infections and person weeks of follow-up was calculated for each cohort in each organisation. This data were analysed using a mixed effects Poisson model, using the natural logarithm of the person weeks as an offset. These are presented in [table 1](#), indicating that there is sufficient power for all but the smallest immune efficacy of 0.1

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%	20%	30%	40%	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

that is, a 10% reduction in incidence in the seropositive cohort. Such a small reduction is indicative of a level of protection unable to provide a means of controlling the pandemic via natural herd immunity.

Estimates for the vaccine effectiveness (VE) element are based on an assumed population of 40 000 participants ([table 2](#)).

The 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 VE estimate 3 months post vaccine, the focus is on those seronegative at baseline.

Estimates for VE 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. Seventy-five per cent 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 (26 000 seronegative of whom 29 500 are vaccinated and 6500 unvaccinated) is assumed.

Statistical analysis plan: primary outcome measure

All enrolled participants will be included in analyses, which will account for clustering by research site. Analyses will be conducted at regular intervals following sufficient events of interest.

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 serostatus 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

Table 2 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 to 74	81 to 95
1	65	46 to 70	85 to 93
2	130	50 to 68	86 to 93
5	325	54 to 65	88 to 92

assessment of the role of factors such as age, gender and 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 for 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, multistate and parametric cure rate models incorporating frailty will also be employed. Bayesian approaches to cure rate models with frailty as describe by de Souza *et al*²⁷ will be employed.

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 Jackson²⁸ will be employed.

Statistical Analysis plan: Vaccine Effectiveness

Survival analysis will be used to estimate the HR in vaccinated compared with 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 VE will be stratified by vaccine manufacturer. VE will also be stratified by baseline positivity (either PCR or antibody), age group (<50, >=50) and time since vaccination (3 month intervals and as a spline). Interaction with sex, ethnicity and risk group will be tested and, if significant, VE 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 VE 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. Testing frequency by arm (positive/negative cohort, vaccination status) will be reported in analyses of the primary outcome. The PCR test for virus is being used as a diagnostic test and hence has high performance. Sufficient sera will be obtained to rerun 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.

Data will be analysed using STATA V.15 and R software.

Study oversight

Oversight is provided by the study management group, chaired by the chief investigator, with representatives from UKHSA, Public Health Scotland, Public Health Wales, Public Health Agency (Northern Ireland) and the COVID-19 Genomics Consortium UK.

The study follow-up period will end by default 24 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 UKHSA.

ETHICS AND DISSEMINATION

The study has received approval from Berkshire Research Ethics Committee on 22 May 2020¹² and has also received support from NIHR as an urgent public health study, which allows central research network resources to recruit participants.

Participants provide informed consent prior to entry to the study and have the option to withdraw at any time. At withdrawal, participants can choose to have their data or samples retained or destroyed, or partial variations. Protocol deviations and breaches will be recorded by the site research teams and the Sponsor will be informed of any serious breaches within one working day.

Dissemination of key study findings and results will take place through regular reports to national expert

committees, preprints, publication in peer-reviewed journals and international conferences. Furthermore, regular communication of study results will be communicated to study sites and participants through national webinars and newsletters. Annotated code for SIREN analyses will be made available at: (<https://github.com/SIREN-study/SARS-CoV-2-Immunity>). The metadata for published SIREN analyses will be available to researchers through the Health Data Research UK CO-CONNECT platform and available for secondary analysis.

DISCUSSION

Strengths

As far as the Authors are aware this study is the largest national longitudinal study examining the question of reinfection with SARS-CoV-2 globally. In a system where staff members may be tested in different settings depending on the timing and reasons for testing (community testing hubs, other hospitals, primary care), the automated method of data extraction and access to national testing data means that the study is less likely to miss potential cases. As far as possible the study is designed to run alongside normal laboratory processes; laboratories use the same assays and procedures which are in place for all other testing, reducing additional burden on sites. The flexible design of this cohort study has enabled the study to adapt to examine VE as vaccines were rolled out in the UK.

The study design reduces bias because we are likely to capture both symptomatic and asymptomatic cases due to frequent participant testing as part of the study, participants are likely to have additional testing through healthcare worker testing programmes (cases identified in this way will be included in the study data) and access to national testing data for cases tested outside the healthcare organisation. The study population is also highly exposed, giving greater power to detect reinfections.

Valuable data are collected on EOI, such as reinfections, including serum samples preceding the EOI which are critical for assessing correlates of protection. Rich serological data are collected in the SIREN biobank, with serial serum samples collected on all individuals with previous infection and vaccination, to permit large-scale longitudinal serological analyses.

The study design lends itself to forming sub-cohorts for more detailed investigations. It has active research collaborations with immunology researchers from the UK Research and Innovation (UKRI) Immunology consortium to investigate T cell responses and with the Wellcome Trust funded Humoral Immune Correlates of COVID-19 consortium to investigate humoral immune responses.

Weaknesses

Cohort retention is an important consideration for the study team, to avoid losing power to detect the primary outcome and potential introduction of bias if there is

differential attrition by cohort. To mitigate this, the study team actively monitor withdrawals and participant feedback, to implement improvements and provide direct participant communications (eg, a newsletter and participant webinars) to promote engagement. Differences in demographics, general health and ongoing risk of exposure between healthcare workers and the general population mean that the results may not be fully generalisable to the UK population.

Author affiliations

¹National Infection Service, UK Health Security Agency, London, UK

²Statistics, Modelling and Economics Unit, UK Health Security Agency, London, UK

³MRC Biostatistics Unit, Cambridge, UK

Acknowledgements The authors would like to thank our partners in the Devolved Administrations, particularly Dr Lesley Price from Glasgow Caledonian University & Public Health Scotland and Dr Muhammad Sartaj, Public Health Agency Northern Ireland for their contribution and advice in preparing the August 2020 protocol amendment.

Contributors SH is the chief investigator and conceived the study. SH, CSB and MCh drafted the first study protocol, with substantial design contribution from MR, MZ and TJGB. AC wrote the statistics plan and power calculations and NE wrote the vaccine effectiveness statistics plan with input from VH and SF. SH and VH are responsible for study delivery, including designing and updating aspects of the study design and submitting protocol amendments. SW, MCo, PDK, NG, AA, JT, SF, AT-K, JI, MS, SR, BO and AV were responsible for designing or updating aspects of the study design and drafting protocol amendments. SW, VH & SH wrote the first draft of the manuscript, with contributions from AC, SF, BO, AV, SR, MS, PDK, MZ, TJGB, MR, CSB, MCh. VH, SW and SH were responsible for updating the manuscript to reflect the current protocol, with contributions as follow: AC and NA were responsible for updating the section on the statistics plan, MCo was responsible for updating the laboratory sections, AA and AT-K for drafting the section on Events of Interest, SF and PDK for updating the section on data management, JI for updating the sections on collaborations, dissemination and participant engagement, SF, JT, NG for updating the sections on recruitment and extended follow-up. The manuscript was submitted after review by and approval from all authors.

Funding This work was supported by the Department of Health and Social Care and UK Health Security Agency (formerly Public Health England), with contributions from the Scottish, Welsh and Northern Irish governments.

Competing interests MZ has unremunerated positions as ISIRV Chair, Member of NERVTAG, SAGE and JCVI and codirector NIHR HPRU, Imperial College London. TB receives funding from the UKRI for the SIREN study. MR's team in the UKHSA Immunisation Department provides vaccine manufacturers (including Pfizer) with postmarketing surveillance reports about pneumococcal and meningococcal disease which the companies are required to submit to the UK Licensing authority in compliance with their Risk Management Strategy. A cost recovery charge is made for these reports. Other authors have no competing interests to declare.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Sarah Wallace <http://orcid.org/0000-0003-1053-2521>

Andre Charlett <http://orcid.org/0000-0001-7154-0432>

Peter D Kirwan <http://orcid.org/0000-0001-6904-0500>

Nick Andrews <http://orcid.org/0000-0003-2069-2684>

Tim J G Brooks <http://orcid.org/0000-0002-3783-1284>

Susan Hopkins <http://orcid.org/0000-0001-5179-5702>

REFERENCES

- Government of the Hong Kong Special Administrative Region. CHP closely monitors cluster of pneumonia cases on mainland 31/12/2019. Available: <https://www.info.gov.hk/gia/general/201912/31/P2019123100667.htm> [Accessed 12 Jan 2022].
- World Health Organization Regional Office for Europe. Who announces COVID-19 outbreak a pandemic, 2020. Available: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic> [Accessed 12 Jan 2022].
- World Health Organization. Who coronavirus (COVID-19) Dashboard. Available: <https://covid19.who.int/> [Accessed 13 Feb 2022].
- Van Elslande J, Vermeersch P, Vandervoort K. Symptomatic SARS-CoV-2 reinfection by a phylogenetically distinct strain. *Clin Infect Dis* 2021;73:354–6.
- To KK-W, Hung IF-N, Ip JD, *et al.* Coronavirus disease 2019 (COVID-19) re-infection by a phylogenetically distinct severe acute respiratory syndrome coronavirus 2 strain confirmed by whole genome sequencing. *Clin Infect Dis* 2021;73:e2946–51.
- Prado-Vivar B, Becerra-Wong M, Guadalupe JJ, *et al.* A case of SARS-CoV-2 reinfection in Ecuador. *Lancet Infect Dis* 2021;21:e142.
- Larson D, Brodnyak SL, Voegtly LJ, *et al.* A case of early reinfection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Clin Infect Dis* 2021;73:e2827–8.
- Tillett RL, Sevinsky JR, Hartley PD, *et al.* Genomic evidence for reinfection with SARS-CoV-2: a case study. *Lancet Infect Dis* 2021;21:52–8.
- Gupta V, Bhoyar RC, Jain A. Asymptomatic reinfection in two healthcare workers from India with genetically distinct SARS-CoV-2. *Clin Infect Dis* 2021;73:e2823–5.
- Atti A, Ferrari M, Castillo-Olivares J, *et al.* Serological profile of first SARS-CoV-2 reinfection cases detected within the siren study. *J Infect* 2022;84:248–288.
- Hall VJ, Foulkes S, Charlett A, *et al.* SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (siren). *Lancet* 2021;397:1459–69.
- Hansen CH, Michlmayr D, Gubbels SM, *et al.* Assessment of protection against reinfection with SARS-CoV-2 among 4 million PCR-tested individuals in Denmark in 2020: a population-level observational study. *Lancet* 2021;397:1204–12.
- Abu-Raddad LJ, Chemaitelly H, Coyle P, *et al.* SARS-CoV-2 antibody-positivity protects against reinfection for at least seven months with 95% efficacy. *EClinicalMedicine* 2021;35:100861.
- Sheehan MM, Reddy AJ, Rothberg MB. Reinfection rates among patients who previously tested positive for coronavirus disease 2019: a retrospective cohort study. *Clin Infect Dis* 2021;73:1882–6.
- Hall V, Foulkes S, Insalata F. Effectiveness and durability of protection against future SARS-CoV-2 infection conferred by COVID-19 vaccination and previous infection; findings from the UK siren prospective cohort study of healthcare workers March 2020 to September 2021. *MedRxiv* 2021.
- Andrews N, Tessier E, Stowe J. Vaccine effectiveness and duration of protection of Comirnaty, Vaxzevria and Spikevax against mild and severe COVID-19 in the UK. *MedRxiv* 2021.
- Aldridge RW, Yavlinsky A, Nguyen V. Waning of SARS-CoV-2 antibodies targeting the spike protein in individuals post second dose of ChAdOx1 and BNT162b2 COVID-19 vaccines and risk of breakthrough infections: analysis of the virus Watch community cohort. *MedRxiv* 2021.
- Seow J, Graham C, Merrick B, *et al.* Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat Microbiol* 2020;5:1598–607.
- Levin EG, Lustig Y, Cohen C, *et al.* Waning immune humoral response to BNT162b2 Covid-19 vaccine over 6 months. *N Engl J Med* 2021;385:e84.
- World Health Organization Headquarters. Enhancing response to omicron SARS-CoV-2 variant: technical brief and priority actions for member states, 2022. Available: [https://www.who.int/publications/m/item/enhancing-readiness-for-omicron-\(b.1.1.529\)-technical-brief-and-priority-actions-for-member-states](https://www.who.int/publications/m/item/enhancing-readiness-for-omicron-(b.1.1.529)-technical-brief-and-priority-actions-for-member-states) [Accessed 13 Jan 2022].
- Chvatal-Medina M, Mendez-Cortina Y, Patiño PJ, *et al.* Antibody responses in COVID-19: a review. *Front Immunol* 2021;12:633184.
- Shields A, Faustini SE, Perez-Toledo M, *et al.* SARS-CoV-2 seroprevalence and asymptomatic viral carriage in healthcare workers: a cross-sectional study. *Thorax* 2020;75:1089–94.
- Grant JJ, Wilmore SMS, McCann NS, *et al.* Seroprevalence of SARS-CoV-2 antibodies in healthcare workers at a London NHS trust. *Infect Control Hosp Epidemiol* 2021;42:212–4.
- Houlihan CF, Vora N, Byrne T, *et al.* Pandemic peak SARS-CoV-2 infection and seroconversion rates in London frontline health-care workers. *Lancet* 2020;396:e6–7.
- Nguyen LH, Drew DA, Graham MS, *et al.* Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. *Lancet Public Health* 2020;5:e475–83.
- Public Health England, University of Oxford, Oxford University Hospitals. Evaluation of sensitivity and specificity of four commercially available SARS-CoV-2 antibody immunoassays. London, 2020. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/898437/Evaluation_of_sensitivity_and_specificity_of_4_commercially_available_SARS-CoV-2_antibody_immunoassays.pdf [Accessed 13 Jan 2022].
- de Souza D, Cancho VG, Rodrigues J, *et al.* Bayesian cure rate models induced by frailty in survival analysis. *Stat Methods Med Res* 2017;26:2011–28.
- Jackson CH. Multi-State Models for Panel Data: The msm Package for R. *J Stat Softw* 2011;38:1–28.