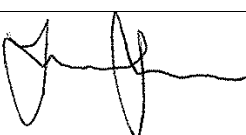# VIVALDI: Statistical Analysis Plan

Full Project Title	Understanding SARS-CoV-2 infection, immunity and its duration in care home staff and residents in the UK
Study registration number	TBC
Version Number	1.2 [First publicly available version]
Previous Versions, nature of changes made	N/A
Study protocol registration	ISRCTN14447421 05/06/2020
Effective Date	18th August 2020
Authors	Dr Andrew Copas, Tom Palmer, Dr Maria Krutikov, Dr Laura Shallcross

Approvals

Name & Title(s)	Signature	Date
Dr Andrew Copas, Statistical Oversight		18/08/2020
Dr Laura Shallcross, Principle investigator		18/08/2020

Contents

Table of Contents

1.   Introduction

1.1 General introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused a large number of deaths worldwide and rapidly changed how people live their lives by restricting social contact and daily activities. Early evidence pointed to the disproportionate impact of coronavirus disease 2019 (COVID-19) on the elderly, ethnic minorities and people with co-morbidities[1,2]. Care home residents have underlying risk factors for severe outcomes (age, comorbidity), but are also likely to have high rates of exposure to infection, through contact with care home staff or other residents. Studies from the USA, Canada and Europe have consistently shown high prevalence of SARS-CoV-2 in care home residents and staff, associated with significant mortality in residents[3–5].

In England an estimated 450,000 individuals aged > 65 years live in approximately 9,000 care homes[6]. Mortality data from the Office of National Statistics (ONS) suggests that >45,000 care home residents have died during the pandemic, although only 12,500 of these deaths were explicitly linked to COVID-19[7]. Accurate estimates of the burden of SARS-CoV-2 infection in care home residents and staff and the proportion of cases without symptoms are lacking because there has to date been limited testing for infection (antibody and RNA tests), and there is no comprehensive surveillance system for infection in care homes. We also have little insight into how infection transmits in the care home, both between staff and residents, and between care homes and other settings (community, hospitals), or how the pandemic has impacted on ways of working in the care home. The prevalence and duration of immunity to SARS-CoV-2 among staff and residents is also unknown.

The Department of Health & Social Care is currently rolling out infection (PCR) testing to all care home staff and residents[8]. This will provide accurate data on the prevalence of infection across all care homes and insights into the types of care homes that are most likely to develop outbreaks. But this large-scale approach is not well-suited to assessing how outbreaks progress over time, or duration of immunity – information that is essential to inform the approach to testing in care homes for current and future pandemic waves.

These questions can be answered most efficiently through a large prospective cohort study of care home staff and residents with repeat testing for infection and analysis of the immune response (antibody and cellular immunity). This will be combined with detailed data collection on symptoms and risk factors for infection, and linkage to NHS and public health data sources

including viral sequencing. Through linkage to an existing study (CATCH-19)[9], we will also undertake qualitative interviews with 30 healthcare workers in a subset of care homes to gain insights into how the pandemic has impacted on healthcare staff and ways of working in the care home. This study is one of the largest undertaken in care homes, and will inform planning and the national public health response to COVID-19

This VIVALDI-2 study will be performed in conjunction with the VIVALDI-1 study. VIVALDI-1 is a cross-sectional survey of all 9000 care homes in England which provide dementia care or care to residents aged >65 years. The study is a collaboration between the ONS, Public Health England and UCL. The survey was delivered to care home managers and collected information on care home characteristics and staffing (number and type of staff, number of residents, sickness pay, use of agency staff), disease control measures (cohorting, isolation, care home closures) and cases of infection in staff and residents. Care home level responses will be linked to individual-level PCR results from the national surveillance whole care home testing data and analysed. Results from the survey are publicly available on the ONS website and have informed government policy on COVID-19 testing in care homes[10].

1.2 Guarantor of the statistical analysis

Dr Andrew Copas

1.3 Staff responsible for performing the statistical analyses

Dr Andrew Copas

Tom Palmer

2. Aims

2.1 General aims of the study and scope of analysis plan

To investigate the epidemiology of SARS-CoV2 infection amongst care home residents and staff in England using intensive testing, through collaboration with Four Seasons Health Care (FSHC), a national chain of care homes.

This analysis plan covers the statistical quantitative analyses for the main results of the study, which will mainly address prevalence of SARS-CoV-2 infection, antibody response and its duration in care home staff and residents in the UK. Separate plans may be prepared with a focus on further topics, such as whole genome sequencing of virological samples.

2.2 Research Questions

2.3 Primary

- What proportion of care home staff and residents have an antibody response to SARS-CoV-2 (i.e. have been previously infected) at baseline, 6 weeks and 3 months?
- How does this vary by care home characteristics (such as size, number of outbreaks, CQC rating) and individual-level characteristics (age, gender, ethnicity)?

2.4 Secondary

1. What proportion of residents and care home staff have acute SARS-CoV-2 infection at each round of swab testing? What proportion of those infections are asymptomatic?
2. What is the incidence of new SARS-CoV-2 infections in residents and staff in the 2-week period following each round of swab testing? How does this differ between care homes with and without an active outbreak (i.e. ≥1 PCR confirmed SARS-CoV-2 cases)?
3. What is the incidence of antibody response between baseline and 3 months among staff and residents with a negative antibody test at baseline? What proportion of those with acute SARS-CoV-2 infection at any round of swab testing do not achieve an antibody respose by 3 months?
4. What is the duration of antibody response in residents who have a response between baseline and 3 months (inclusive)?
5. What are the quantitative levels of IgG and IgM in those with an antibody response between baseline and 3 months, and how do the levels change over time from the first response?
6. What is the direct and indirect mortality in care home residents attributable to SARS-CoV-2?

3. Study Design

3.1 Study population

Staff and residents who live or work in all FSHC residential or nursing homes in England (N=105) are eligible to participate in the study. FSHC will provide the study team with a list of eligible care homes and the name and telephone details for each care home manager. In order for the care home to be enrolled, the care home manager must confirm their willingness

to participate.   We aim to include all residents aged > 65 years and all staff in the study, the estimated numbers eligible are 6,500 staff and 5,000 residents.

Informed consent to participate will be collected in writing from all care home staff. Wherever possible, written informed consent will be sought directly from residents. Study information materials will be translated into other languages upon request.   For residents who lack capacity, the care home will contact a friend or next of kin to act as personal consultee. If a personal consultee cannot be identified, a nominated consultee, such as the care home manager or a member of staff will be asked to complete a professional declaration form. This approach meets the four criteria for HRA approval as set out in sections 31-33 of the Mental Capacity Act 2005[11]. This approach is also justified by the current lack of research in this population that can address important questions regarding impact and transmission of COVID-19.

3.2 Study design

The study is a prospective cohort study, enrolling participants in May 2020, with follow-up ending in April 2021. The study aims to capture both the current and anticipated second wave of the pandemic.   Intensive testing will be carried out for all enrolled participants. All participants will undergo testing with nasopharyngeal swabs to detect SARS-CoV-2 viral RNA using PCR testing and blood testing to detect antibodies to this virus and cellular immunity. Two further rounds of PCR testing will be repeated at 7-28 day intervals with accompanying symptom data capture. A blood sample for antibody testing will also be undertaken for all enrolled staff and residents at approximately 6 weeks and 3 months. Further antibody testing in a subset of residents who have evidence of an antibody response at 3 months will be undertaken at 6 and 12 months to assess duration of the antibody response. Finally, saliva will be collected from a subset of residents and staff in 4-8 care homes to enable comparison of SARS-CoV-2 detection using salivary versus nasopharyngeal samples.

Table 1. Timing of testing

Nasopharyngeal swab (PCR) testing	Blood sample testing (approximate timings)
1. National testing of all residents and staff in all care homes in England 2. A second round of testing at all participating care homes, 7-28 days after the first testing round 3. A third round of testing at all participating care homes 7-28 days, after the second testing round	1. Baseline testing (all partcipants) 2. 6-week testing (all participants) 3. 3-month testing (all participants) 4. 6-month testing (residents with positive antibody response at 3-months) 5. 12-month testing (residents with positive antibody response at 6-months)

3.3 Participant selection procedures

All eligible residents and staff of participating FSHC residential or nursing homes in England will be considered for enrolment.

3.4 Exclusion/Inclusion criteria

Inclusion criteria:

- All residents aged over 65 years and staff in FSHC residential or nursing homes in England are eligible to participate provided they can speak English

Exclusion criteria:

- None.

3.5 Sample size calculations

Assuming that around 80% of eligible staff and residents will participate, then on average approximately 30 residents and 47 staff will be tested per home across the 105 homes. Based on a conservatively estimated intra-cluster correlation of 0.36 (twice that reported for seasonal influenza)[12] leads to effective sample sizes of around 278 residents and 284 staff, after

accounting for the loss of precision due to clustering. Assuming an antibody prevalence of around 30% for both groups the precision of the estimate (based on 95% confidence interval width) will be ±5.4% for residents and for staff separately.

Importantly, as we anticipate that at least 800 residents will have been infected at 3 months (using a conservative estimate of 20% infection prevalence), we should have sufficient precision and power to detect changes in antibody prevalence over time. For example assuming 80% of infected residents agree to repeated antibody testing and 70% of these have an antibody response at 3 months (450 residents) then we can estimate the proportion of these retaining a response at 6 months within 4.6% assuming a retention rate of 50%. This assumes that antibody response is broadly unaffected by care home level factors.

4. Data and data validation

4.1 Data to be collected

Table 2 summarises the main data collected in the study.

Table 2. Summary of data collected.

Information type:	Measurement:	Data source	When collected
Individual-level (collected for all enrolled participants, unless marked with *)			
Baseline characteristics	• Age • Gender • Date of care home entry/ employment start date • Date of exit	FSHC administrative data	Monthly
Ethnicity	• Ethnicity recorded at time of baseline blood testing for staff and residents	FSHC ethnicity	Baseline
SARS-CoV-2 PCR test results - nasopharyngeal swab	• Date and binary outcome of test for all staff and residents tested since March 1st 2020 (includes some data on who had symptoms)	Foundry - National care home testing dataset	Daily/weekly

Symptoms	• Symptom checklist recorded for all staff and residents on date of swab (PCR) testing
	• Record of which staff or residents develop symptoms in the 6 days following swab testing
	• Symptoms to be checked:
	  o Cough
	  o Shortness of breath
	  o Chest pain
	  o Fever
	  o Loss of taste or smell
	  o Sore throat
	  o Runny nose
	  o Abdominal pain
	  o Diarrhoea
	  o Vomiting
	  o New muscle aches
	  o New headache
	  o New joint pains
	• Number of days with symptoms will also be recorded
		FSHC – data on symptoms	1-3 reports per care home

SARS-CoV-2 PCR test results - nasopharyngeal swab II*	• Date and binary outcome of tests undertaken in care home staff since March 1st 2020 (data will be incomplete)	Foundry – PHE testing dataset	Quarterly

SARS-CoV-2 PCR test results - saliva test*	• Binary result and CT threshold (collected from 4-8 care homes)	Foundry – PCR saliva test results	Once
Antibody test results	• Binary result and antibody titre	Foundry – antibody results	3-5 rounds
Antibody test results validated by Oxford ELISA	• Validation of 20% of all antibody test results set (20% of all samples)	Foundry – Oxford antibody test results	Once
Hospital admission/attendance and reason	• Staff and residents: Date of A&E attendance or hospital admission; date of hospital discharge; primary and secondary ICD-10 diagnostic codes.	Foundry – HES linkage	Quarterly
Death	• Cause of death • Timing of death	Foundry – ONS mortality	Quarterly

Aggregate-level (collected for all participating care homes)

Care home baseline characteristics I	• Number of nursing/residential/dementia beds per care home • Geographic location • CQQ rating • Number of staff by role • Staff turnover and new starters (TBC)	FSHC administrative data	Baseline
Care home baseline characteristics II	• Number of staff and roles • Staff/resident contact frequency • Staff sick pay (yes/no) • Dates of closure to visitors • Dates of closure to admissions • Use of bank/agency staff • Staff sickness • COVID-19 cases / hospital admissions/deaths (residents) • Use of disease control measures • PPE supply	National care home survey – May/June 2020	Baseline
Care home building characteristics	• Layout of care home • Description of building	FSHC care home survey	Baseline
Care home number of visitors	• Daily visitor count and type since 1st March 2020 (TBC)	FSHC care home survey	Quarterly

| Care home COVID-19 reports | • Number of suspected and confirmed COVID-19 cases
• Hospital admissions (residents)
• Deaths (residents)
• Staff absences | FSHC daily reports | Daily |

Acronyms: A&E: accident and emergency, CT: cycle threshold, FSHC: Four Seasons Health Care, HES: Hospital Episode Statistics, ICD-10: International Statistical Classification of Diseases and Related Health Problems 10th Revision, ONS: Office for National Statistics, PCR: polymerase chain reaction PHE: Public Health England, PPE: personal protective equipment.

4.2 Assurance of data quality

To be considered later as this is primarily the responsibility of NHS Foundry.

4.3 Consistency and error checking

To be considered later as this is primarily the responsibility of NHS Foundry.

4.4 Missing data and evidence of selection bias

It is uncear how much missing data (principally missing tests) will occur among residents and staff whilst they remain at the home but it is certainly expected that some residents (and staff) may leave due to transfer, hospitalisation or death. A flow chart will be produced to capture the numbers of participants moving in and out of the participating homes over time, and the numbers accepting and refusing tests. Descriptive analysis will be conducted to investigate any differences between care homes who were invited but declined participation and included care homes. A descriptive analysis of potential selection bias will also be conducted in all care homes, by comparing eligible staff and residents that consent to participate in the study with those who did not consent, by age and gender. Although individual-level data is not expected to be available for those declining to participate, using summary statistics or age and gender amongst those participating in combination with aggregate care-home level data for these variables will enable estimation of the average age and proportion female of those not consenting in each care home.

Since there may be intermittent missingness in the antibody testing data (i.e. some individuals at the home tested on some occasions but not on others), we later consider how this can be handled if it occurs (see section 6.7) either through imputation or (for time-to-event analyses) by censoring.

4.5 Constructed variables

- Symptomatic. A binary indicator measuring whether participant reported any COVID-19 symptoms. This will also be disaggregated into binary indicators for:
    - At least one common COVID-19 symptom (Cough, Shortness of breath, Chest pain, Fever, Loss of taste or smell)
    - Only less common COVID-19 symptoms (Sore throat, Runny nose, Abdominal pain, Diarrhoea, Vomiting, New muscle aches, New headache, New joint pains)
    - Summary statistics will be provided separately for all measured symptoms.

  This will be done separately for symptoms at time of testing and cumulatively for symptoms reported over the following six days. This will enable the separate construction of variables for asymptomatic, pre-symptomatic and symptomatic.

- Incidence of antibody response at 3 months among staff and residents with a negative antibody test at baseline. A binary indicator will be constructed, taking the value of one if an individual had a negative antibody response at baseline and a positive antibody response at either 6 weeks or 3 months.

- Indicator of acute infection during study period. A binary indicator will be constructed, taking the value of one if an invidiual tested positive for acute infection at any of the rounds of swab testing conducted as part of this study.

- Time-varying indicator of infection history. A time-varying categorical variable that indicates whether a participant had no infection history, had a prior infection with no antibody response, or had a prior infection with antibody response at the most recent available testing round.

5.   Outcome measurement and summary statistics

5.1 Primary outcome

The primary outcome of the study is measured antibody response to SARS-CoV-2 at baseline, 6 weeks and 3 months. Antibody test results will be categorised as a binary outcome using immunity thresholds that have been established in the tests initial evaluation[13]. A positive test result suggests that the individual tested has previously been infected with SARS-CoV-2 at the time of the antibody testing. This outcome will be summarized to give an estimate of the prevalence of antibody response amongst all participants. All participants will receive antibody testing at baseline, 6 weeks and 3 months. The primary outcome at baseline is defined simply by the baseline antibody response.

However at 6 weeks and 3 months we define the primary outcome in two different ways for analysis: as a positive result at that time point, and cumulatively by any positive test result up to and including that time point.

Provided levels of missed tests are low, then any missed tests are considered negative when defining the 'cumulative version' of the primary outcome, though participants who miss all three tests are excluded. In the event of more substantial missingness (e.g >10% of tests among participants with 1+ tests conducted) then imputation will be used (see section 6.7) to impute the test result. In considering missing data we do not consider a test result missing if the participant has died or moved away from the home on a long-term basis and no imputation will be used for these participants.

5.2 Secondary outcomes

- *Acute infection at the time of each round of swab testing.* A binary indicator reflecting whether the individual tested has a current infection. This will be measured in 1-2 rounds of testing at 7-14 intervals, depending on whether the care home has an active outbreak (i.e. ≥1 PCR confirmed SARS-CoV-2 cases). Infections will be also be summarized by the presence of symptoms (see section 4.5). This outcome will be summarised after each testing round to estimate the proportion of residents and care home staff who are infected with SARS-CoV-2.
- *New antibody response to SARS-CoV-2 at 3 months.* A binary indicator of whether a participant who tested negative for antibody response at baseline has a positive antibody test at either 6 weeks or 3 months.
- *Antibody response to SARS-CoV-2 at 6 and 12 months.* This will be measured in a subset of residents who have evidence of an antibody response at 3 months. This will inform assessment of the duration of antibody response. Antibody testing at 6 and 12 months will not be undertaken in staff since the issue of duration of the antibody response in healthcare workers is being addressed through other cohorts.
- *Quantitative IgG and IgM reponse to SARS-CoV-2.* This will be summarised by the median and IQR calculated for 'positive' antibody tests at baseline, 6 weeks, 3 months (all staff and residents), and 6 and 12 months (residents with antibody response at 3 months)
- Hospital admission. Linkage to Hospital Episodes Statistics data for hospital admissions will enable estimation of the impact of SARS-CoV-2 on hospital admission (direct and indirect).

- Mortality. Linkage to ONS mortality will enable estimation of the impact of infection and the antibody response on death (direct and indirect). Estimate the direct and indirect mortality in care home residents attributable to SARS-CoV-2
- Consent and non-consent to the study are recorded in the recruitment log. The overall % consenting will be reported.
- Participants are free to withdraw from the study at any time. This will be outlined in the participant information sheets and in the consent forms. They can do this by talking to the care home manager or by contacting the study manager.

6. Statistical Analyses

6.1 Primary outcome

For the primary outcome we will estimate separately the proportion of staff and residents who have an antibody response to *SARS-CoV-2* at baseline, 6 weeks and 3 months. Confidence intervals for these proportions will be presented, based on robust standard errors to acknowledge the clustering within care homes. Antibody test results will be categorised as a binary outcome using immunity thresholds that have been established in the tests initial evaluation (5). In staff and residents we will report the proportion with antibodies but no prior record of infection testing or symptoms (residents) or sickness absence (staff). These estimates will be repeated for each round of antibody testing, based on both the 'current' and 'cumulative' definitions of the primary outcome (see earlier). Summary statistics will also be presented for prevalence amongst staff and residents at the care home level, we will report the proportion of care homes with no evidence of prior infection, and provide a figure of the % of staff and residents with antibody response across homes.

Additionally, univariate analysis will be conducted using response and care home characteristics (such as size, number of outbreaks, CQC rating) and individual-level characteristics (such as ethnicity, age and gender) to identify variables associated with a positive antibody test. Statistically significant variables will then be included in a multivariate logistic regression, with random effects for care home.

The baseline data will be analysed separately and reported first (prevalence and associations), and then the analysis will be repeated based on 3 months data, with the

outcome defined as described earlier. No analysis of associations with the data at 6 weeks will be conducted.

In the event of the antibody tests having <95% sensitivity or specificity when validated by the Oxford Elisa then we will also report an estimated prevalence of antibodies at baseline, 3 and 6 months corrected for test performance.

6.2 Secondary outcomes

Acute infection at the time of each round of swab testing

The proportion of staff and residents (separately) who are infected in each round of PCR testing (based on RT-PCR of nasopharyngeal swabs) will be reported, together with the proportion of infected individuals that are asymptomatic. Confidence intervals will take account of clustering by care home as outlined above. We will also estimate an incidence rate of new infection, based on data collected during the intensive testing period from particants who are negative at their first test. We will undertake risk factor analysis for infection, at either first or subsequent testing round or detected subsequently. Potential risk factors include both care home characteristics (such as size, number of outbreaks, CQC rating) and individual-level characteristics (such as age, gender, ethnicity). Risk factor analysis will be based on logistic regression with random effects for care homes. The incidence rate of new infection will be presented separately by the number of suspected or confirmed cases in the care home at the previous round of testing, in order to compare the rates of infection in care homes with established outbreaks and undetected outbreaks.

New antibody response to SARS-CoV-2 by 3 months.

Amongst staff and residents with a negative antibody test at baseline, we will report the proportion of staff and residents who subsequently develop an antibody response as measured by antibody testing at 6 weeks and 3 months. As above, confidence intervals for these proportions will account for clustering within care homes with robust standard errors. We will also separately present the proportion developing an antibody response amongst those participants testing positive for acute infection at any of the rounds of swab testing.

The incidence rate of new response will also be calculated. As testing could be intermittent we propose in this rate estimation to censor participants at 6 weeks with a negative result at 6 weeks and no test at 3 months, and to consider those testing positive at 3 months without a test at 6 weeks to have developed their response at 6 weeks (midpoint from baseline).

Participants with missing tests will be assumed not to have gained and lost an antibody response within 3 months, so a negative test at 3 months and missing test at 6 weeks is interpreted as negative throughout. In calculating the rate, death and moving away are regarded as censoring events, so the rate estimated can be seen as the rate given the time at risk (i.e. whilst inviduals are alive and at risk of devloping a response).

The same risk factors considered for the primary outcome will also be considered for this secondary outcome. A Poisson regression will be conducted taking account of 'time at risk', and including random frailty terms for care homes.

Duration of antibody response

The estimates described above for the primary outcome will also be repeated for 6- and 12-month rounds of antibody testing in a subset of residents with a positive antibody response at 3 months. We will first describe duration of the antibody response by reporting the proportion of individuals tested with an antibody response (based on thresholds outlined above) at 6 and 12 months. We will also conduct a 'time-to-event' analysis using test results from baseline to 12 months, setting 'time zero' to the time of the first positive antibody test (noting that first antibody response will predate the first test result in each case). In this analysis participants will be censored at their last negative test, and the timing of the event is not recorded in continuous time but is subject to interval censoring (e.g. known to have occurred between 3 and 6 months). We will plot a Kaplan-Meier curve and conduct Cox regression including predictors, first individually and then jointly, as outlined for the analysis of the primary outcome with random frailty terms for care homes.

Quantitatve IgG and IgM trajectory

We shall report the median and IQR at each round of testing for positive test results (including 6 and 12 months performed only for those with antibody response at 3 months. We shall also report the median and IQR on the 'time scale' of time from first antibody response, to be interpreted alongside the analysis described above of time to loss of response.

Hospital admissions

Linkage to data on the secondary outcome of hospital admissions will enable estimation of the impact of infection and the antibody response on the number of admissions. In order to allow for multiple hospital admissions, this will be analysed using a Poisson regression model with random frailty terms for care homes. Predictive categorical variables will be included in the

model, including a time-varying indicator of infection history (i.e. never infected, prior infection with no antibody response at last test, and prior infection with antibody response at last test), in addition to control variables such as age and gender. Given the frequency of testing, and to ensure antibody status remains recent, we initially propose analysing admissions in all participants over the first 6 months from baseline (i.e. assuming that antibody status at 3 months remains relevant at 6 months). However, this assumption will be reviewed based on emerging data on the duration of antibody response. For example, we may analyse admissions for all particpants only over the first 4 months if the data suggests a shorter duration of antibody response.

We shall also repeat the analysis focussing on the impact of antibody retention, analysing admissions between 3 and 12 months after baseline, and restricted to participants with an antibody response at 3 months. In this analysis we will consider any participants with no antibody response at 6 months to have lost their response at 4.5 months, and those first antibody negative at the 12 month test to have lost their response at 9 months. Again, these assumptions will be reviewed based on the duration of antibody response observed in the data.

Mortality

Linkage to mortality data will enable estimation of the impact of infection and the antibody response on death (direct and indirect). We shall follow the same approach as for analysis of admissions using a time varying predictor of infection and antibody status (i.e. never infected, prior infection with no antibody response at last test and prior infection with antibody response at last test). However, in this case a Cox regression model will be used.

6.3 Interim analysis

Some of the above analysis, including baseline estimates of the primary outcome, will be conducted after baseline testing. Analysis will take place on a continuous basis throughout the duration of the study, as data from each round of testing becomes available.

6.4 Sub-group analysis

None beyond what is already described. Note all analysis is conducted separately for care home staff and residents, and there is no formal testing of difference between them.

6.5 Corrections for multiple testing

Considered to not be required.

6.6 Sensitivity analysis

To be considered.

6.7 Imputation

This will be considered if there is an appreciable missingness from an intermittent pattern of antibody testing performed over time for participants, whilst they remain living or working at the care home. If conducted then multiple imputation based on chained equations will be used, using observed test results to predict missing ones, and with separate modelling within each care home (subject to sufficient care home size).

6.8 Other analysis

Statistical analysis plans for viral sequencing data and comparison with Virus Watch will be developed in due course, since these datasets will not be available for several months. A qualitative sub-study will also be developed to assess the extent to which test results inform disease control measures, and whether this improves disease control

6.9 Statistical software

Stata® version 15.

References

1. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* **395**, 1054–1062 (2020).

2. Lewer, D. *et al.* Black Asia and Minority Ethnic groups are at increased risk from COVID-19: indirect standardisation of NHS mortality data. *Wellcome Open Res.* **5**, 88 (2020).

3. McMichael, T. M. *et al.* Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. *N. Engl. J. Med.* (2020) doi:10.1056/NEJMoa2005412.

4. Observer Reporters. Across the world, figures reveal horrific toll of care home deaths.

The Guardian.

5.   Fisman, D., Lapointe-Shaw, L., Bogoch, I., McCready, J. & Tuite, A. Failing our Most Vulnerable: COVID-19 and Long-Term Care Facilities in Ontario. *medRxiv* 2020.04.14.20065557 (2020) doi:10.1101/2020.04.14.20065557.

6.   LaingBuisson. *Care homes for older people UK market report*. (2019).

7.   Office for National Statistics. *Deaths involving COVID-19 in the care sector, England and Wales: deaths occurring up to 1 May 2020 and registered up to 9 May 2020 (provisional)*.
      https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/deathsinvolvingcovid19inthecaresectorenglandandwales/deathsoccurring upto1may2020andregisteredupto9may2020provisional#:~:text=Of   deaths   involving COVID-19 among ca (2020).

8.   Department of Health & Social Care. *Coronavirus (COVID-19) Scaling up our testing programme*.
      https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/878121/coronavirus-covid-19-testing-strategy.pdf (2020).

9.   CATCH-19. *COVID-19: Impact and Burden in Care Homes*. Accessed August 18, 2020. https://pass-antibiotics.github.io/catch19/ (2020).

10.   Office for National Statistics, University College London, Department of Health & Social Care, Public Health England & Ipsos Mori. *Impact of coronavirus in care homes in England: 26   May   to   19   June   2020*.
      https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/impactofcoronavirusincarehomesinenglandvivaldi/26mayto19june2020#collaboration (2020).

11.   Department for Constitutional Affairs. *Mental Capacity Act 2005*. Section 31-33 (2005).

12.   Hayward, A. C. et al. Comparative community burden and severity of seasonal and pandemic influenza: Results of the Flu Watch cohort study. Lancet Respir. Med. (2014) doi:10.1016/S2213-2600(14)70034-7.

13.   Adams, E.R., Ainsworth, M. & Anand, R. Antibody testing for COVID-19: A report from the National COVID Scientific Advisory Panel. *medRxiv* (2020)