04.03.20

Serological Surveillance of COVID-19 in England: Sera Collection Protocol

1. Background and rationale

Severity assessment of COVID-19 requires the ability to detect asymptomatic and mild infections and thus determine the true number of infections within the general population. The number of true infections within the population provides the denominator for the estimation of severity measures such as infection -fatality and infection -hospitalisation ratios. The number of true infections can be determined if the prevalence of immunity of the population prior to, during and after the epidemic are known. This information is provided by serological surveys.

Currently serological methods to diagnose infection are being developed and once validated these assays can be used to provide valuable information to understand the extent of transmission of SARS-CoV-2 and monitor how this changes over time. If this data is available early, it can be used to adjust planning assumptions and help predict the impact upon health and social care services.

The UK undertook a series of influenza sero-surveys during the 2009 pandemic based upon residual blood samples from the HPA National Sero-epidemiology Programme supplemented by samples from chemical pathology labs (1-3). These samples were submitted to HPA and NHS laboratories for routine diagnostic purposes at that time. Although this work delivered critical information on background population seroprevalence and sero-incidence, several issues were raised in post-pandemic reviews. These reviews highlighted that, although this information was gathered and published earlier than almost any other country, even earlier availability of this intelligence would have been critical to inform important national policy decisions. The following key recommendations have been made in relation to flu sero-epidemiology and could be applied to the COVID pandemic:

- The Science and Technology Committee (3rd report 2010-11) (4) stated that sero-epidemiological data needs to be available **earlier** in the time course of a future pandemic to help with risk assessment;
- SPI (May 2011 06(07)(01)), JCVI flu sub-committee and the JCVI highlighted the importance of **establishing routine sero-surveillance** for seasonal influenza;
- The CMO-SLG determined that **sero-surveillance is critical to determine population immunity and community infection rates. This data cannot be obtained from other sources and is vital to making modelling predictions of the pandemic**.
- The 2011 UK Influenza Pandemic Preparedness Strategy includes **sero-epidemiology as a key surveillance initiative that will be required at the start of any pandemic, and states that work should be underway to enhance capability in this area based on the H1N1** (2009) influenza pandemic (5) .
- The Hine Report states that **further exploration of population-based surveillance**, such as serology, should be considered (6).

- The SPI-M group recommended that the HPA **strengthen population-based influenza sero-epidemiology,** including collection of key epidemiological information on vaccination status and underlying risk status.
- The PHE influenza programme has identified the development of sampling strategies and methods for establishing representative population surveys as an R and D priority;
- Finally the European Centre for Disease Control and Prevention (**ECDC**) has highlighted the importance of influenza sero-epidemiology, as has **WHO** as part of the Fineberg report (7) into the pandemic response has highlighted the need for a proper assessment of severity at national and subnational levels early in a pandemic. **Sero-epidemiology forms a key element of that severity assessment.**

As a result of these recommendations a range of population-based serum banks have been used by PHE for monitoring influenza and other diseases across England.

2. Aim

To set out the requirements to measure the sero-prevalence of SARS-CoV-19 in England during 2020.

3. Surveillance methodology

The project proposes to undertake a population-based seroprevalence survey, involving several thousand individuals across the age-range using sera from before and during the COVID-19 outbreak.

Serological analysis of appropriately age stratified and geographically representative samples will provide information on seroprevalence and sero-incidence to set in context other measures of impact of COVID-19 in the UK population, and provide the most accurate measures of population exposure.

The work will consist of three interlinked work-packages: 1) Population survey 2) Laboratory analysis 3) Statistical and modelling analysis.

4. Sampling

4.1 Sampling frame

Sero-surveys will be carried out using sera collected and submitted to the PHE Sero-epidemiology Unit (SEU) through three approaches:

1. Residual sera collected in secondary care and submitted to the PHE Sero-epidemiology Unit (SEU). The PHE Sero-epidemiology Unit (SEU) archive is an opportunistic collection of residual serum samples from routine microbiological testing, submitted voluntarily each year from laboratories throughout England. In 2019, Manchester, Newcastle, Exeter, Leicester laboratories participated in this voluntary collection, with approximately 400 samples received per month. It is anticipated that these

laboratories would continue to participate in 2020, however the number of samples expected will be negatively affected due to the COVID-19 situation. In 2009, chemical pathology laboratories were recruited to contribute to SEU collections and use of this route will be investigated for this study. SEU archive sera are stored at the PHE North West regional laboratory in Manchester and are anonymised and permanently unlinked from any patient identifying information, with only age, gender, date of collection (if available) and contributing laboratory retained.

2. Residual sera collected from approximately 100 GP practices across the country participating in RCGP Research and Surveillance Centre (RSC) primarily for Influenza surveillance. This collection of samples is opportunistic, with samples collected from individuals requiring routine blood testing via their GP and provides a geographically representative dataset for England.
3. What's The Story COVID-19 will build on an existing national research network and ethically-approved NIHR-funded study to collect childhood and teenage serum samples for near real-time monitoring of increases in paediatric COVID-19 sero-positivity rates across the UK in 2020. Sample Size will be 100 per month in each of the 0-4, 5-9, 10-14 and 15-19 year-old age-groups, with at least 200 samples per age group taken before any widespread transmission (i.e. baseline). The challenge of obtaining blood samples from representative cohorts of healthy children had already led to the Oxford Vaccine Group, Public Health England and regional partners throughout England establishing the 'What's the STORY (Serum Testing of Representative Youngsters)' network to evaluate antibody levels against vaccine-preventable diseases. This study has commenced, and with additional funding would be expanded to collect 3200 samples throughout 2020, including collection of individual level information on recent respiratory illnesses and relevant medical history. Further expansion would be possible as needed. The region partners are as follows: Oxford, South London/Kent, Yorkshire and Humber, Bristol, Southampton, & Manchester. Postcodes have been selected to be representative of the population according to index of multiple deprivation, although the expanded COVID-19 recruitment will aim to achieve sample size targets by additional advertising for individuals to volunteer to participate in the areas.

4.2 Sample representativeness

For population susceptibility, a total of 2,000 existing sera collected prior to the start of circulation of COVID-19 will be required. This will then be followed by monthly sero-surveys with 1000 samples per month to estimate sero-incidence. The complete age range will be represented. The sample will be stratified with 200 for pre-pandemic survey (100 for monthly) specimens in the following age groups (1-4, 5-9, 10-14, 15-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70+).

4.3 Sample size and precision

The precision (95%CI) of incidence estimates based on simple subtraction of prevalence, based on approximately 1,000 samples, collected from 8 age groups, are shown in Tables 1a-b. Note that in practice incidence and cumulative prevalence will be modelled statistically over time which will mean greater precision that that shown. For prevalence the precision based on 100 and 200 samples in each age group is shown in table 1c.

Table 1: Incidence percentages and 95% CI for different number of samples per age group

(a) Comparison to baseline from a specific month

Prevalence at specific month (n=100)	Baseline prevalence (n=200)		
	0%	5%	10%
5%	5 (0.7,9.3)		
10%	10 (4.1,15.9)	5 (-1.6,11.6)	
15%	15 (8.0,22.0)	10 (2.4,17.6)	5 (-3.1,13.1)
20%	20 (12.2,27.8)	15 (6.6,23.4)	10 (1.1,18.9)
30%	30 (21.0,39.0)	25 (15.5,34.5)	20 (10.1,29.9)
40%	40 (30.4,49.6)	35 (24.9,45.1)	30 (19.5,40.5)
50%	50 (40.2,59.8)	45 (34.7,55.3)	40 (29.4,50.6)

(b) Comparison between consecutive months for monthly incidence

Prevalence the following month (n=100)	Prevalence in one month (n=100)					
	0%	5%	10%	15%	20%	25%
5%	5 (0.7,9.3)					
10%	10 (4.1,15.9)	5 (-2.3,12.3)				
15%	15 (8.0,22.0)	10 (1.8,18.2)	5 (-4.1,14.1)			
20%		15 (6.1,23.9)	10 (0.2,19.8)	5 (-5.5,15.5)		
25%			15 (4.7,25.3)	10 (-1,21)	5 (-6.6,16.6)	
30%				15 (3.6,26.4)	10 (-1.9,21.9)	5 (-7.4,17.4)
35%					15 (2.8,27.2)	10 (-2.6,22.6)
40%						15 (2.2,27.8)

(c) precision of prevalence estimates (95%CI) for n=100 and n=200

Prevalence	N=100	N=200
0	0.0,3.6	0.0,1.8
5	1.6,11.3	2.4,9.0
10	4.9.17.6	16.2,15.0
15	8.6,23.5	10.4,20.7
20	12.7,29.2	14.7,26.2
25	16.9,34.6	19.2,31.6
30	21.2,40.0	23.7,36.9
40	30.3,50,3	33.2,47.1
50	39.8,60.2	42.9,57.1

Note that the sample size calculations shown are based on a binary serological classification of positive or negative based on an appropriate cut-point.

5. Sample analysis

Serology samples will be analysed at the Respiratory Virus Unit (RVU) at PHE Colindale. Each sample will be tested once and not in duplicate. Approximately 1-2ml of serum is required to perform a combination of assays:

- 1st line: a screening ELISA assay using whole COVID-19 virus lysate (currently undergoing validation).
- 2nd: a spike (S1) specific ELISA (currently in development). The S1 portion of the spike protein contains the receptor binding side of the virus. We therefore assume that antibodies detected in this assay can be used as a proxy for functional antibody – this link needs to be established by measuring neutralising antibody.
- 3rd: a confirmatory test – performed by neutralisation using UK isolate virus in cell culture (currently in development).

6. Statistics and modelling

6.1 Descriptive analysis

A cut off for seropositivity will be determined following assay development and testing.

Descriptive statistics will describe the population sampled, numbers of samples and give the sero-prevalence and sero-incidence estimates (with 95%CI) for samples taken in each month by age group and by region. A weighted overall estimate will also be calculated allowing for the sampling fractions by age and region. This will also enable estimation of numbers of infection when combined with ONS population data. In addition, geometric mean titres and reverse cumulative distribution (RCD) curves will be calculated with 95% CI's.

6.2 Statistical modelling of cumulative incidence

Appropriate statistical models of temporal trends in sero-positivity across age groups and regions would be fitted to the cumulating data (cumulative incidence), e.g. fractional polynomial logistic regression. This would need to be appropriately weighted to obtain population level seroprevalence estimates, particularly when samples from "What's the STORY" are included. Note that modelled estimates of sero-positivity would be somewhat more precise than those presented in Table 2 and calculated in the descriptive analysis.

6.3 Models of epidemic spread

Models for real time prediction of epidemic spread have been developed during 2009 (Birrell et al, 2011, Baguelin et al, 2011). Serological data is fundamental to ensuring that modelled estimates of the epidemic are correctly scaled to provide a reliable estimate of the numbers of infections occurring in the population. These data provided information on the "baseline" age-specific susceptibility before the epidemic. Also, in conjunction with other surveillance data collected during the epidemic and indirectly related to incidence of infection (e.g. GP consultations, positivity data), these data have been shown to be crucial in the reconstruction of the incidence of infection by age group and region. These models will be used together with future available data on different aspects of the disease (e.g. hospitalisations) to provide real time estimates, of infections and forecasts of the likely future course of the epidemic. The ability of these models to track incidence of infection in a timely fashion and to provide reliable projections depends crucially on the prompt availability of valid serological data, which, as the epidemic evolves, informs the scale of the pandemic. We, therefore, aim to establish a system of rapid handling of routine serological data, including data extraction, transfer and analysis.

7. Proposed outputs

- A serosurveillance pandemic COVID-19 data-set use in real-time models.
- Temporal estimates of population susceptibility and sero-incidence due to COVID-19
- A report or publication to disseminate findings.

8. Funding and resources

Table 3

Sampling			
Sample Source	**Actions**	**Estimated cost**	**Organisation responsible for funding**
SEU	Agreed increase in funding allocated for laboratories to send samples to SEU from £4 to £8.	??	? Is this increase PHE funded
RCGP	Agreed to extend surveillance programme to cover COVID-19 sampling	Funding outside PHE	RCGP
What's the STORY expansion	Bid put in to NIHR	Funding outside PHE if bit accepted	Oxford Vaccine Group / NIHR
Testing			
Component	**Actions**	**Estimated cost**	**Organisation responsible for funding**
Test assay	Assay in development	TBC	? PHE
Associated lab costs of running of testing		??	PHE
Staffing			

Staff	WTE requirement	Estimated cost	Organisation responsible for funding
Project lead			PHE
Laboratory scientist			PHE
Surveillance			PHE

9. Ethical approval

a. SEU - PHE has ethical approval (05/Q0505/45) for the collection and use of unlinked and anonymised residual serum samples in cross-sectional antibody prevalence studies for the surveillance of population immunity to vaccine preventable diseases of public health importance and the collection has been extensively used for this purpose.

b. RCGP RSC - this work is Health Protection and therefore falls under Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002 the existing agreement with PHE. Therefore, no further legal basis is required. The implications of this are that verbal consent (as now for virology samples) is all that is required for virological and serological samples. This has been confirmed by email with PHE's Caldicott Guardian's Office.

c. COVID-19: Coronavirus STORY - use of collected serum for COVID-19 sero-epidemiology is possible under the ethical approval for 'What's the STORY' (NRESXXXXXXXX), in parallel to submission of substantial amendments to increase the study sample size and modify recruitments methods to enable an even and sustained recruitment rate.

10. Timelines and milestones

Key milestones for the work are highlighted below and the associated timelines in Table 3.

Table 4: Project timelines

Month 0	Month 1	Month 2-7	Month 8
▪ Convene Steering Group ▪ Initiate request for sample collection	▪ Collect and analyse residual sera pre-season	▪ Collect and analyse residual sera monthly ▪ Undertaken statistical and mathematical analysis	▪ Undertake final monthly sero-survey

References

1. Hardelid P, Andrews N, Hoschler K, Stanford E, Baguelin M, Waight P, et al. Assessment of baseline age-specific antibody prevalence and incidence of infection to novel influenza A/H1N1 2009. Health technology assessment (Winchester, England). 2010;14(55):115-92.
2. Baguelin M, Hoschler K, Stanford E, Waight P, Hardelid P, Andrews N, et al. Age-specific incidence of A/H1N1 2009 influenza infection in England from sequential antibody prevalence data using likelihood-based estimation. PLoS one. 2011;6(2).
3. Miller E, Hoschler K, Hardelid P, Stanford E, Andrews N, Zambon M. Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study. The Lancet. 2010;375(9720):1100-8.
4. Committee TSaT. Scientific advice and evidence in emergencies: third report of session 2010-11, Vol. 1: Report, together with formal minutes, oral and written evidence: The Stationery Office; 2011.
5. Team DPIP. UK Influenza Pandemic Preparedness Strategy 2011. 2011 2011.
6. Hine D. The 2009 influenza pandemic: an independent review of the UK response to the 2009 influenza pandemic: Cabinet Office; 2011.
7. Organization WH. Strengthening response to pandemics and other public-health emergencies: report of the review committee on the functioning of the International Health Regulations (2005) and on pandemic influenza (H1N1) 2009. 2011.