Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection

Version: 1.1
Date: 17 March 2020
Contact: EarlyInvestigations-2019-nCoV@who.int

Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection

Protocol summary

Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection	
Study population	Age-stratified convenience sample from general population
Potential output and analysis	Estimates of: • Extent of age-specific infection • Cumulative incidence of infection • Infection attack rates • Fraction of asymptomatic infection • Case fatality ratio
Study design	Prospective population-based convenience sample from the general population, stratified by age.
Study duration	The investigation can be conducted as a cross-sectional investigation, or can include serial sampling as a prospective cohort study.
Minimum information and specimens to be obtained from participants	Data collection: Epidemiological data including: clinical symptoms. Specimens: Serum samples to inform seroepidemiological inferences.

WHO, in collaboration with technical partners, has developed a series of enhanced surveillance protocols, that are harmonized to help provide detailed insight into the epidemiological characteristics of COVID-19. Other COVID-19 investigations and studies protocols currently available include:

- Household transmission investigation protocol for COVID-19 virus infection
- Protocol for assessment of potential risk factors for COVID-19 infection among health care workers in a health care setting.
- Surface sampling of COVID-19 virus: A practical "how to" protocol for health care and public health professionals.

All WHO protocols for COVID-19 are available on the WHO website together with the technical guidance documents. This currently includes case definitions, laboratory guidance, infection prevention and control and travel advice.

1 Background

The detection and spread of an emerging respiratory pathogen are accompanied by uncertainty over the key epidemiological and serologic characteristics of the novel pathogen and particularly its ability to spread in the human population and its virulence (case-severity). This is the case for the COVID-19 virus, first detected in Wuhan city, China in December 2019 (1).

To date initial surveillance has focused primarily on patients with severe disease, and, as such, the full spectrum of the disease, including the extent and fraction of mild or asymptomatic infections that do not require medical attention are not clear. Estimates of the case fatality ratio, and other epidemiological parameters, will likely be lower than current crude mortality estimates once the full spectrum of disease is able to be included in the denominator. In addition, the role of asymptomatic or subclinical infections in human-to-human transmission of COVID-19 virus is not well understood and it is not yet clear whether those who are reported as asymptomatic may be able to transmit the virus to other individuals.

With a novel coronavirus, initial seroprevalence in the population is assumed to be negligible due to the virus being novel in origin. Therefore, surveillance of antibody seropositivity in a population can allow inferences to be made about the extent of infection and about the cumulative incidence of infection in the population.

The following protocol has been designed to investigate the extent of infection, as determined by seropositivity in the general population, in any country in which COVID-19 virus infection has been reported. Each country may need to tailor some aspects of this protocol to align with public health, laboratory and clinical systems, according to capacity, availability of resources and cultural appropriateness. However, using a standardized protocol such as this one below, epidemiological exposure data and biological samples can be systematically collected and shared rapidly in a format that can be easily aggregated, tabulated and analyzed across many different settings globally for timely estimates of COVID-19 virus infection severity and attack rates, as well as to inform public health responses and policy decisions. This is particularly important in the context of a novel respiratory pathogen, such as COVID-19 virus.

Comments for the user's consideration are provided in purple text throughout the document as the user may need to modify methods slightly because of the local context in which this study will be carried out.

1.1 Objectives

There are two primary objectives for this seroepidemiological investigation:

1. To determine the extent of infection in the general population and age-specific infection cumulative incidence, as determined by seropositivity
2. To determine the fraction of asymptomatic or subclinical infections

Seroepidemiological investigations provide the opportunity to evaluate secondary objectives, such as, but not limited to:

1. To determine risk factors for infection by comparing the exposures of infected and non-infected individuals
2. To more accurately determine the case fatality ratio

COMMENT: Little is currently known about COVID-19 virus antibody kinetics. Asymptomatic infected persons may clear the virus more quickly than do symptomatic patients and antibody titers in the former are likely to be lower, if they seroconvert at all, than in infected patients exhibiting symptoms. These are considerations for the interpretation of any COVID-19 virus seroepidemiological investigation.

2 Study procedures

2.1 Study design

The seroepidemiological investigation for COVID-19 virus infection is a population-based, age-stratified prospective study. It is intended to provide key epidemiological and serologic characteristics of COVID-19 virus.

There are three possibilities for how this study can be implemented:

1) Cross-sectional investigation
2) Repeated cross-sectional investigation in the same geographic area (but not necessarily the same individuals each time).
3) Longitudinal cohort study with serial sampling of the same individuals each time

COMMENT: The first option will likely be the easiest to implement for countries to implement, while the third provides the most comprehensive information on extent of infection, as described below. The choice as to how this study will be implemented should be determined by feasibility and available capacity.

The timing of the study will depend on the specific public health questions that need to be addressed. If serial sampling is to be conducted, it is best to initiate the investigation as quickly as possible. Serial sampling can then be conducted as long as possible, as determined by capacity and resources.

For one-time cross-sectional investigations, there may be an interest in completing the investigation after the peak of transmission of the epidemic wave. However, a cross-sectional investigation, conducted at any time of the epidemic, will provide important information that can be used to inform public health responses.

For longitudinal cohort study with serial sampling, the epidemic curve from surveillance (daily number of new confirmed cases) can be used to adjust the frequency with which samples are collected to provide real-time estimates of seropositivity in the general population.

2.2 Study population

The geographic scope of the investigation should first be defined. This may be limited to a local or regional investigation, or may be conducted as a national investigation. Within the geographic scope of the study, high incidence and low incidence areas should be identified. The selection of these areas should be informed by the latest information on COVID-19 virus circulation, available on the WHO website.

The study population should then be identified in at least one high incidence and one low incidence area through the random selection of households. For the purpose of this investigation, a household is defined as a group of people (2 or more) living in the same residence. In practice, the technical definition may vary due to social, political and cultural practices.

Definitions of a household which may be used (but are not limited to):

- Two or more people living together in a domestic residence (residential institutions, such as boarding schools, dormitories, hostels or prisons will be excluded).
- A dwelling or group of dwellings with a shared kitchen or common opening onto a shared household space.

All persons living in the household should be invited to participate in the study, including children to ensure age-specific attack rates can be calculated.

COMMENT: The distribution of household compositions in the population would be needed to calculate population, age-specific infection attack rates.

The random selection of households within a geographic area is one method of recruitment. Random digit dialing[1], blood donors or other ongoing longitudinal cohort studies may also be used to identify and recruit individuals to participate in the investigation. The advantage of working with blood donors is that they are usually forthcoming to being contacted for future follow-up and you may be able to track long-term antibody dynamics. For COVID-19, the age-specific attack rates in blood donors are likely to be similar to that in the general population except for those with substantial comorbidities or elevated exposure (e.g. healthcare workers).

COMMENT: Plasma may be collected from a subset of participants so that the serologic assay results of the serum and plasma could be compared. This might be useful if sera and plasma give the same results as plasma are stored in many blood donor archives which could be used to retrospectively to estimate the infection attack rate.

Whichever method is used to identify and recruit participants, all attempts should be made to include participants over a range of ages in order to be able to determine and compare age-specific attack rates, although crude age-specific estimates will need to be adjusted for age structures in the population.

COMMENT: Depending on which method of study recruitment is chosen, the group implementing the study may choose either to conduct home visits to collect data and specimens or to centralize data and specimen collection at one location, asking participants to travel to the location to participate in the study. Decisions as to how to implement the study should be determined by feasibility and resource (including personnel) availability.

2.3 Eligibility criteria

Inclusion criteria: All individuals identified for recruitment into the investigation, irrespective of age.

Exclusion criteria: Refusal to give informed consent, or contraindication to venipuncture.

COMMENT: Suspected or confirmed acute or prior COVID-19 infection should not be considered as an exclusion criterion for this investigation. Doing so would underestimate the extent of infection in the population. For individuals currently receiving medical care for COVID-19 infection, a family member or proxy may be used to complete the questionnaire on his/her behalf.

[1] Riley S, Kwok KO, Wu KM et al. Epidemiological characteristics of 2009 (H1N1) pandemic influenza based on paired sera from a longitudinal community cohort study. *PLoS Med*. 2011 Jun;8(6):e1000442

2.4 Data collection

Each participant recruited into the investigation should be asked to complete a questionnaire which covers demographic and exposure information. An example of an investigation questionnaire which may be used can be found in the Appendix A. This questionnaire is not exhaustive and may need to be adapted to the local setting and outbreak characteristics, but it provides an outline as to the data to be collected in order to calculate the epidemiological parameters (see 4.3 Epidemiological parameters).

2.5 Specimen collection

A serum sample needs to be collected from each participant upon recruitment into the investigation. The collection of serum samples should follow specimen collection guidance in the country.

Table 1 describes when data and specimens should be collected according to the study design selected. If repeated sampling is to be conducted, whether as a repeated cross-sectional investigation, or as a longitudinal investigation, specimens should be collected at least 21 days apart.

COMMENT: Other specimens (e.g. nasopharyngeal) may be collected to determine acute COVID-19 infection, as determined by the objectives of the investigation and the available resources and capacity.

Figure 1: Schedule for data and specimen collection according to study design

Study design	**Baseline**	**Further recruitment and follow-up (at least 21 days after baseline)**	**Regular follow-up of same individuals recruited (at least 21 days apart)**		
Cross-sectional investigation	Data and specimen collection				
Repeated cross-sectional investigation	Data and specimen collection	Data and specimen collection			
Longitudinal cohort investigation	Data and specimen collection		Data and specimen collection	Data and specimen collection	Data and specimen collection

2.6 Specimen transport

All those involved in the collection and transportation of specimens should be trained in safe handling practices and spill decontamination procedures. For details regarding the transport of samples collected and infection control advice, please refer to case management algorithm and laboratory guidance in the country or WHO laboratory guidance, available on the WHO website.

For each biological sample collected, the time of collection, the conditions for transportation and the time of arrival at the study laboratory will be recorded. Specimens should reach the laboratory as soon as possible after collection. If the specimen is not likely to reach the laboratory within 72 hours, specimens should be frozen, preferably at -80°C, and shipped on dry ice. It is, however, important to

avoid repeated freezing and thawing of specimens. The storage of serum specimens in domestic frost-free freezers should be avoided, owing to their wide temperature fluctuations. Serum should be separated from whole blood and can be stored and shipped at 4°C or frozen to -20°C or lower and shipped on dry ice.

Transport of specimens within national borders should comply with applicable national regulations. International transport of specimens should follow applicable international regulations as described in the WHO Guidance on Regulations for the Transport of Infectious Substances 2019-2020.

2.7 Ethical considerations

Ethical requirements will vary by country. In some countries, this investigation may fall under public health surveillance (emergency response) acts and may not require ethical approval from an Institutional Review Board.

2.7.1 Informed consent

The purpose of the investigation will be explained to all individuals identified for recruitment into the investigation. Informed consent will be obtained from all individuals willing to participate in the investigation before any procedure is performed as part of the investigation by a trained member of the investigation team. Consent for children under the legal age of consent will be obtained from a parent or legal guardian. Each participant must be informed that participation in the investigation is voluntary and that s/he is free to withdraw, without justification, from the investigation at any time without consequences and without affecting professional responsibilities.

COMMENT: The age of consent may vary by country. Check the requirements of local, regional or national authorities.

Informed consent will seek approval to collect blood and epidemiological data for the intended purpose of this investigation, that samples may be shipped outside of the country for additional testing and that samples may be used for future research purposes.

2.7.2 Risks and benefits for subjects

This investigation poses minimal risk to participants, involving the collection of a small amount of blood. The primary benefit of the study is indirect in that data collected will help improve and guide efforts to understand extent of COVID-19 virus infection and may prevent further transmission of the virus.

2.7.3 Confidentiality

Participant confidentiality will be maintained throughout the investigation. All subjects who participate in the investigation will be assigned a study identification number by the investigation team for the labelling of questionnaires and specimens. The link of this identification number to individuals will be maintained by the investigation team and the Ministry of Health (or equivalent) and will not be disclosed elsewhere.

If the data is shared by the implementing organization to WHO or any agency or institution providing support for data analysis, data shared will include only the study identification number and not any personably identifiable information.

Article 45 of the IHR (2005) describes the "treatment of personal data".[2] Person identifiable data collected under the IHR should be kept confidential and processed anonymously, as required by national law. However, such data may be disclosed for assessments and management of public health risks, provided the data are processed fairly and lawfully.

2.7.4 Prevention of COVID-19 virus infection in investigation personnel

All personnel involved in the investigation need to be trained in infection prevention and control procedures (standard contact and droplet precautions, as determined by national or local guidelines). These procedures should include proper hand hygiene and the correct use of surgical masks, if necessary, not only to minimize their own risk of infection when in close contact with individuals with COVID-19 infection, but also to minimize the risk of spread among other participants in the investigation.

WHO technical guidance on infection prevention and control specific to COVID-19 can be found on the WHO website.

3 Laboratory evaluations

Laboratory and biosafety guidance for COVID-19 can be found on the WHO website.

Serologic assays specific to COVID-19 are currently under development / in the process of evaluation. The protocols or SOPs will be published on the WHO website once they become available. Cross reactivity to other coronaviruses may be an issue and should be considered in the interpretation of data. Multiple assays may be required to confirm a seropositive for COVID-19 virus.

Laboratory procedures involving sample manipulation must be carried out in a biosafety cabinet (BSC).

3.1 Serological testing

Serum samples should be screened for the presence of COVID-19 virus specific antibodies using serological testing. Tests for both IgM and IgG should be carried out using an enzyme linked immunosorbent assay (ELISA) or immunofluorescence. If a sample is positive for either IgM or IgG a plaque reduction neutralization test (PRNT) should be done.

ELISA testing should be carried out in a facility with at least biosafety level 2 (BSL-2) capacity.

3.2 Confirmation of the presence of neutralizing antibodies

A PRNT can be carried out on samples that were positive for COVID-19 specific IgM or IgG antibodies to confirm the presence of neutralizing antibodies.

PRNT should be carried out in a facility with at least BSL-3 capacity.

3.3 Sample storage

In the case that serum samples cannot be processed immediately, they should be stored at -80°C. It is recommended to aliquot samples prior to freezing, to minimize freeze thaw cycles.

[2] World Health Organization. International Health Regulations (2005)

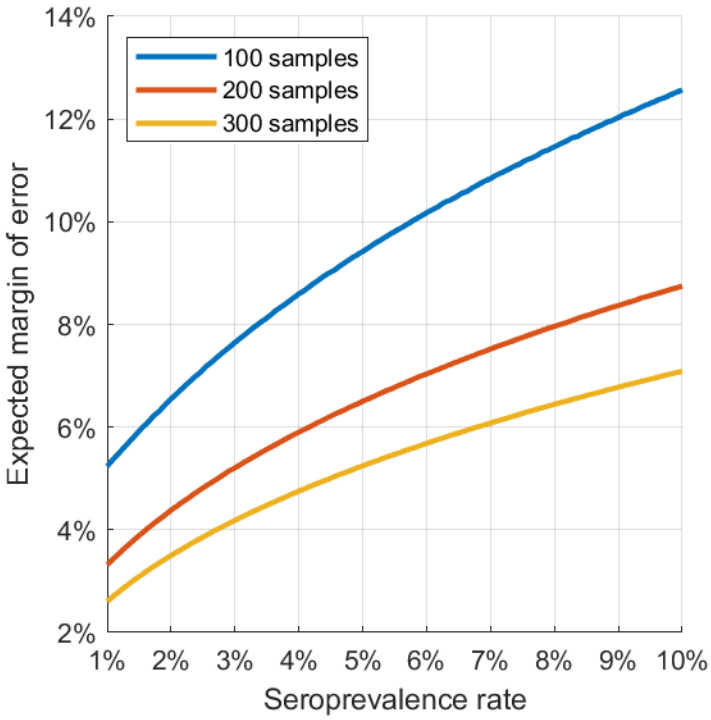
COMMENT: These recommendations are subject to changes as new, reliable serological assays become available.

COMMENT: If serological testing is not available in the country in which serum samples are collected, they may be stored or shipped to an international reference laboratory. WHO is able to facilitate communication with international referral laboratories in order for samples to be shipped for further testing.

4 Statistical analyses

4.1 Sample size

The figure below provides estimates of margin of error as a function of seroprevalence for 100, 200 and 300 samples. For a given seroprevalence rate p and sample size N, the expected margin of error corresponds to the expected width of the 95% confidence interval associated with the point estimate of p obtained using binomial likelihood.



4.2 Epidemiological parameters

The table below provides an overview of the epidemiological parameters that can be measured as part of this investigation.

Parameter	**Definition** (*in bracket*: "simplified" expression of it)	**Form and questions where to get the data to calculate the parameters concerned**	**Comments, limitations**
Age-specific attack rate	The proportion of individuals per age strata who show seropositivity for COVID-19 virus infection		
Age-specific cumulative incidence	The proportion of individuals per age strata who show seropositivity for virus infection		*Proportion should be adjusted for any difference in the age stratification of the participants and the overall population
Symptomatic proportion of cases (asymptomatic fraction)	The proportion of individuals who show symptoms or signs of COVID-19 infection		*The numerators of interest are the numbers of those individuals reporting various signs and symptoms of infection (e.g. fever, cough) and the number/proportion of individuals reporting no signs or symptoms (i.e. the asymptomatic fraction); the denominator is the total number of individuals tested.
Serological response to infection	The change in serum level of specific antibodies to COVID-19 virus (*Increase in titre*)		
Population groups most at risk	The identification of groups who are most vulnerable to COVID-19 virus infection (e.g. age groups, gender, occupation)		*May only be an early signal, a nested case-control study could be conducted to evaluate risk factors for infection
Ratio of severe disease	The proportion of an age group with severe infection, divided by the probability that an infection resulted in a severe case, expressed as a proportion of the total number of people in that age group		
Case fatality ratio	The proportion of individuals with fatal outcome for COVID-19 infection		* May require extended follow-up to determine outcome of those with COVID-19 infection

5 Reporting of findings

5.1 Reporting

Any investigation of this nature should include reporting on the following information:

(1) the number of households and the number of individuals included;
(2) the age and sex of all individuals included
(3) the time in the outbreak of sample collection and the antibody titre levels of each specimen collected.
(4) the number of individuals with serologic evidence of COVID-19 virus infection. If sample size permits, these numbers should be stratified by age.
(5) the number of individuals with serologic evidence of COVID-19 virus infection who have reported symptoms.

It is also important to fully document the study design, how individuals were recruited, and the serological assay and methods used to ensure that data can be pooled to increase power in estimating epidemiological parameters.

Ideally, information would be collected in a standardized format according to the questionnaires and tools in this generic protocol to assist with data harmonization and comparison of results (see forms in Appendix A).

To enable results to be aggregated across study sites and across county sites and by extension, strengthen the statistical power of the results, it is expected that they are shared with WHO by sending individual, de-identified data to EarlyInvestigations-2019-nCoV@who.int. The data shared should include only the study identification number and not any personably identifiable information.

6 References

6.1 Further references for COVID-19

WHO Situation reports
https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/

Surveillance and case definitions
https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/surveillance-and-case-definitions

Laboratory guidance
https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance

Clinical management
https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/patient-management

Infection prevention and control
https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/infection-prevention-and-control

Risk communications
https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/risk-communication-and-community-engagement

7 Acknowledgments

This generic protocol was adapted from the protocol entitled “Prospective longitudinal cohort study of influenza infection during epidemic periods” by the Consortium for the Standardisation for Influenza Seroepidemiology (CONSISE). CONSISE is a global partnership aiming to develop influenza investigation protocols and standardise seroepidemiology to inform public health policy for pandemic, zoonotic and seasonal influenza. This international partnership was created out of a need, identified during the 2009 H1N1 pandemic, for better (standardised, validated) seroepidemiological data to estimate infection attack rates and severity of the pandemic virus and to inform policy decisions.

Appendices

Appendix A: Sample questionnaires - Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection

Form 1: Report Form for all participants

Form 2: Laboratory results for all participants

Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection

Form 1: Report Form for all participants

Unique ID	

1. Data Collector Information	
Name of data collector	
Data collector Institution	
Data collector telephone number	
Mobile number	
Email	
Form completion date (DD/MM/YYYY)	___/___/___
Date of interview with informant (DD/MM/YYYY)	___/___/___

2. Identifier information	
First name	
Surname	
Sex	□ Male □ Female □ Not known
Date of Birth (DD/MM/YYYY)	___/___/___
Telephone (mobile) number	
Age (years, months)	
Email	
Country of residence	
Nationality	
Ethnicity (optional)	
Occupation	
Have you had contact with a anyone with suspected or confirmed COVID-19 virus infection?	□ Yes □ No □ Unknown If Yes, dates of last contact (DD/MM/YYYY): ___/___/___

3. Symptom history	
In the past (X) months, have you had any of the following: COMMENT: (X) period to cover time since emergence of COVID-19 virus to date of data collection	
Fever ≥38°C	□ Yes □ No
Chills	□ Yes □ No
Fatigue	□ Yes □ No
Muscle ache (myalgia)	□ Yes □ No
Sore throat	□ Yes □ No
Cough	□ Yes □ No
Runny nose (rhinorrea)	□ Yes □ No
Shortness of breath (dyspnea)	□ Yes □ No
Wheezing	□ Yes □ No
Chest pain	□ Yes □ No
Other respiratory symptoms	□ Yes □ No
Headache	□ Yes □ No
Nausea/vomiting	□ Yes □ No

Abdominal pain	□ Yes □ No
Diarrhoea	□ Yes □ No
Did any of these symptoms require you to seek medical attention?	□ Yes □ No □ Unknown
Did any of these symptoms require you to miss work or school?	□ Yes □ No □ Unknown
Did any of these symptoms require you to be hospitalized?	□ Yes □ No □ Unknown

Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection

Form 2: Laboratory results

This table will need to be completed for every serum sample collected, as determined by the chosen specimen collection schedule and design of the study.

19b. Serology testing methods and results (complete new table for each specimen collected):	
Lab identification number	
Date sample collected (DD/MM/YYYY)	(DD/MM/YYYY)___/___/___
Date sample received (DD/MM/YYYY)	(DD/MM/YYYY)___/___/___
Type of sample	□ Serum □ Other, specify:
Type of serological assay	
Serology result	□ Positive □ Negative □ Unknown
COVID-19 virus titres	
Date of result (DD/MM/YYYY)	___/___/___
Specimen shipped to other laboratory for confirmation - Date (DD/MM/YYYY)	□ Yes □ No (DD/MM/YYYY)___/___/___
Confirmation result	□ Positive □ Negative □ Unknown