Population-based age-stratified seroepidemiological investigation protocol for coronavirus 2019 (COVID-19) infection

Version: 2.0
Date: 26 May 2020
Contact: EarlyInvestigations-2019-nCoV@who.int

World Health Organization

Reference:

The emergence of a new virus means that understanding transmission patterns, severity, clinical features and risk factors for infection will be limited at the start of an outbreak. To address these unknowns, WHO has provided Four Early sero-epidemiological Investigation Protocols (rebranded the WHO Unity Studies). One additional study to evaluate environmental contamination of COVID-19 is also provided.

These protocols are designed to rapidly and systematically collect and share data in a format that facilitates aggregation, tabulation and analysis across different settings globally.

Data collected using these investigation protocols will be critical to refine recommendations for case definitions and surveillance, characterize key epidemiological features of COVID-19, help understand spread, severity, spectrum of disease, and impact on the community and to inform guidance for application of countermeasures such as case isolation and contact tracing.

They are available on WHO website here: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/early-investigations)

COVID-19 investigations and studies protocols currently available include:

1. The First Few X cases and contacts (FFX) investigation protocol for coronavirus disease 2019 (COVID-19).

2. Household transmission investigation protocol for coronavirus disease 2019 (COVID-19)

3. Protocol for assessment of potential risk factors for coronavirus disease 2019 (COVID-19) among health workers in a health-care setting.

4. Population-based age-stratified seroepidemiological investigation protocol for coronavirus 2019 (COVID-19) infection

5. Surface sampling of COVID-19 virus: a practical "how to" protocol for health care and public health professionals

Please contact earlyinvestigations-2019-nCoV@who.int for any questions.

All WHO protocols for COVID-19 are available on the WHO website together with the technical guidance documents.

Version Control

Main updates to version 2.0 by each section include:

- (Objectives): More specific objective 1 which is to measure sero-prevalence in the population by age group. Inclusion of a further secondary objective to contribute to an improved understanding of antibody kinetics following COVID-19 infection.
- (Recruitment of population): Ideal age groupings for which age-specific sero-prevalence should be reported are proposed. Greater detail provided of options for convenience sampling are described and less for general household surveys.
- (Specimen collection): deletion of Figure 1.
- (Serological testing): Inclusion of rapid diagnostic tests (RDTs) as an option for immunoassays to be employed in sero-epidemiological studies.
- (Confirmation of presence of neutralizing antibodies): Confirmation procedure to include equivocal results. Confirmation procedures can be performed on a sample.
- (Sample size): Reference to open-source on-line sample size calculator. Sample size estimates are likely to increase for household surveys as a consequence of design effect. For serial sampling investigations, sample size calculations need to be powered to detect differences between the different rounds of the survey.
- (Epidemiological indicators): Indicators described in Table 1 fully updated and related to specific objectives of a sero-epidemiological investigation.
- (Reporting): Minimal reporting should also include study design and response rates.
- (Questionnaires) Appendices: addition of symptoms, update of the laboratory reporting form and also changes to the suggested reporting forms to ensure standardization with other Unity studies protocols.
- Alignment of structure and format to technically edited other Unity studies protocols

Contents

Summary

Population-based age-stratified seroepidemiological investigation protocol for COVID-19 infection	
Study population	Age-stratified convenience sample from general population
Potential output and analysis	Estimate or inform estimates of: • Seroprevalence of antibodies to COVID-19 • 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 There are three possibilities for how this study can be implemented: 1) One-time cross-sectional investigation 2) Repeated cross-sectional investigation in the same geographic area (but not necessarily sampling the same individuals each time) 3) Longitudinal cohort investigation with serial sampling of the same individuals
Study duration	The study can be conducted as a "one time "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 basic demographics and clinical symptoms Specimens: Serum samples to inform seroepidemiological inferences

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. Background

1.1 Introduction

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 transmissibility (i.e. ability to spread in a population) and its virulence (i.e. case-severity). This is the case for the SARS-CoV-2 virus, first detected in Wuhan city, China in December 2019.

To date initial surveillance has focused primarily on patients with symptoms or 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 estimates once the full spectrum of disease is able to be included in the denominator. In addition, the role of pre-symptomatic, asymptomatic or subclinical infections in human-to-human transmission of SARS-CoV-2 virus is not well understood.

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-19s.

1.2 Objectives

There are **two primary objectives** for this sero-epidemiological investigation:

1. To measure the **seroprevalence of antibodies to COVID-19 in the general population by sex and age group,** in order to ascertain the cumulative population immunity; and
2. To estimate the **fraction of asymptomatic, pre-symptomatic or subclinical infections in the population and by sex and age group.**

Sero-epidemiological investigations provide the opportunity to inform or evaluate **secondary objectives**, such as, but not limited to:

3. To determine **risk factors for infection** by comparing the exposures of infected and non-infected individuals;
4. To contribute to determine the **case fatality ratio**; and
5. To contribute to an improved understanding of **antibody kinetics** following COVID-19 infection.

COMMENT: Little is currently known about COVID-19 virus antibody kinetics. Asymptomatic infected persons may clear the virus more quickly than do symptomatic patients. Antibody titers in the asymptomatic persons 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 sero-epidemiological investigation.

2 Methods

2.1 Design

The sero-epidemiological 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 SARS-CoV-2 virus.

There are three study designs that can be used:

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

COMMENT: The first option will likely be the easiest for countries to implement, while the third provides the most comprehensive information on attack rates, as described below. The choice as to how this study will be implemented should be determined by the objectives, feasibility and available capacity (e.g. capital, financial, and personnel).

The **timing of the study** will depend on the specific public health questions that need to be addressed.

- *One-time cross-sectional investigations*: there may be an interest in completing the investigation after the first or subsequent peaks of transmission of the epidemic waves. 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.
- *Repeated cross-sectional and longitudinal cohort investigations* both entail serial sampling, either from different or the same individuals. It is best to initiate these investigations as early as possible. Serial sampling can then be conducted as long as possible, as determined by capacity and resources. Intervals between each round of collecting specimens should be of a period of greater than 21 days. For longitudinal cohort investigations 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 Population

The geographic scope of the investigation should be defined. This may be limited to a **local** or **regional** investigation, or may be conducted as a **national** investigation.
Ideally, the geographic scope of the investigation should be **representative of the overall burden of infection (i.e. include both high and low incidence areas)** and this selection can be informed by the latest information on SARS-CoV2 virus circulation, available on the WHO website.

2.3 Recruitment of population

The method to recruit the investigation population will depend on the objectives, the feasibility and the resources available to conduct the study.

Whichever method is used to identify and recruit the investigation population, all attempts should be made to include participants **over a range of ages** in order to determine and compare age-specific sero-prevalence. Crude age-specific estimates will need to be adjusted for age structures in the population. Ideally, investigations should ensure that the following 10 age groups can be reported: 1-4, 5-9, 10-14, 15-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70+.
COMMENT: If age-groups detailed above are not feasible, investigators should ensure that age categories employed are coherent with these. Reporting age-specific indicators for the younger ages by 5-year age bands (i.e. 1-4, 5-9, 10-14, and 15-19) will better inform plans for loosening lock down policies (i.e. the opening of schools). The younger groups can be collapsed into 10-year age bands (i.e. 1-9 and 10-19) if this is more feasible.

Populations can be recruited **by using existing studies** (e.g. general health surveys, research population, or patient cohorts, etc.) **or by establishing new studies**.

To recruit participants, **sampling** can be either by:

- ***Convenience Sampling*:** Individuals attending medical services (e.g. blood donors, pregnant mothers, primary care attendees, etc.) can be approached to participate in the study. 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). However, some age-groups will not be captured.
  Convenience samples can also be constructed using residual sera taken from patients for other investigations. Investigations using residual sera can be easier to implement and can reflect the exposure in the general population, but the information collected can be limited (e.g. location, age and sex).
- ***Random Sampling***: Individuals can be approached to participate in investigations using a variety of probability sampling techniques such as random digit dialing[1] and general household surveys (e.g. in low middle income countries Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), population-based HIV impact assessment (PHIA). Households are often defined as a group of people (2 or more) living in the same residence, but in practice, the technical definition may vary due to social, political and cultural practices. It usually excludes residential institutions, such as boarding schools,