Nationwide population-based surveys of Iranian COVID-19 Serological Surveillance (ICS) program: The surveys protocol

Kazem Khalagi[1], Safoora Gharibzadeh[2], Davood Khalili[3,4], Siamak Mirab Samiee[5], Seyed Mahmoud Hashemi[6], Saeide Aghamohamadi[7], Maryam Mir-Mohammad-Ali Roodaki[8], Katayoun Tayeri[9], Hengameh Namdari Tabar[10], Kayhan Azadmanesh[11], Jafar Sadegh Tabrizi[12], Kazem Mohammad[13], Samira Goudarzi[14], Firoozeh Hajipour[15], Saeid Namaki[6], Alireza Raeisi[16], Afshin Ostovar[1]* [ID]

Abstract

Background: Serological surveillance of COVID-19 through conducting repetitive population-based surveys can be useful in estimating and monitoring changes in the prevalence of infection across the country. This paper presents the protocol of nationwide population-based surveys of the Iranian COVID-19 Serological Surveillance (ICS) program.

Methods: The target population of the surveys is all individuals ≥6 years in Iran. Stratified random sampling will be used to select participants from those registered in the primary health care electronic record systems in Iran. The strata are the 31 provinces of the country, in which sampling will be done through simple random sampling. The sample size is estimated 858 individuals for each province (except for Tehran province, which is 2574) at the first survey. It will be recalculated for the next surveys based on the findings of the first survey. The participants will be invited by the community health workers to the safe blood sampling centers at the district level. After obtaining written informed consent, 10 mL of venous blood will be taken from the participants. The blood samples will be transferred to selected reference laboratories in order to test IgG and IgM antibodies against COVID-19 using an Iranian SARS-CoV-2 ELISA Kit (Pishtaz Teb). A serologically positive test is defined as a positive IgG, IgM, or both. After adjusting for the measurement error of the laboratory test, nonresponse bias, and sampling design, the prevalence of COVID-19 will be estimated at the provincial and national levels. Also, the approximate incidence rate of infection will be calculated based on the data of both consecutive surveys.

Conclusion: The implementation of these surveys will provide a comprehensive and clear picture of the magnitude of COVID-19 infection and its trend over time for health policymakers at the national and subnational levels.

Keywords: COVID-19, Serological, Surveillance, Sero-prevalence, Antibody tests, Survey, Nationwide, Population-based, Iran

_____

Corresponding author: Dr Afshin Ostovar, aostovar@tums.ac.ir

1. Osteoporosis Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran
2. Department of Epidemiology and Biostatistics, Research Centre for Emerging and Reemerging Infectious Diseases, Pasteur Institute of Iran, Tehran, Iran
3. Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
4. Department of Biostatistics and Epidemiology, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
5. Reference Health Laboratory, Ministry of Health and Medical Education, Tehran, Iran
6. Departments of Immunology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
7. Deputy of Health, Ministry of Health and Medical Education, Tehran, Iran
8. Reference Health Laboratory, Ministry of Health and Medical Education, Tehran, Iran
9. National HIV/AIDS care and treatment advisor, Center for Communicable Disease Control, Ministry of Health and Medical Education, Tehran, Iran
10. HIV/STI Control Department, Center for Communicable Disease Control, Ministry of Health and Medical Education, Tehran, Iran
11. Virology Deportment, Pasteur Institute of Iran, Tehran, Iran
12. Tabriz Health Services Management Research Center, Iranian Center of Excellence in Health Management, Tabriz University of Medical Sciences, Tabriz, Iran
13. Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
14. Department of Epidemiology and Biostatistics, Pasteur Institute of Iran, Tehran, Iran
15. Endocrinology Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran
16. School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran

↑*What is "already known" in this topic:*
So far, several sero-prevalence studies have been performed in some metropolises or provinces of Iran or on high-risk populations. The management of COVID-19 outbreak at the province and country level requires reliable estimates of the prevalence of infection in the general population, which has not been provided through the results of the previous studies.

→*What this article adds:*
The regular nationwide population-based serological surveys for COVID-19 in Iran will provide the information needed to manage the epidemic at the national and subnational levels.

Introduction

Since the outbreak of COVID-19 in Wuhan, China, in December 2019, the disease has spread rapidly around the world and has become one of the major health problems in recent decades (1). In Iran, since the official report of the first definite case on February 19, 2020, more than 1,305,399 definite cases and more than 56,457 deaths due to this disease have been reported until January 13, 2021.

Currently, the definitive method of diagnosing COVID-19 is the use of genomic techniques, such as PCR, and sequencing performed using swab sampling from the throat or nasopharynx. In this method, as in other diagnostic methods, cases of false negatives are observed, which are related to reasons such as the time interval from the onset of infection and the sampling bugs. In the case of many known pathogenic viruses, serological findings have been a common method for examining the history of pathogen exposure. Compared to PCR, serological testing has always been desired due to the shorter time to achieve the result and less cost and workload. Serological tests are used to measure the response of antibodies to the virus and to monitor the severity of disease transmission and social vulnerability (2, 3).

Since COVID-19 epidemiological surveillance through PCR is often performed for symptomatic people, the use of serological tests are important in epidemiological surveillance of the epidemic to identify people exposed to the virus and developed antibodies against it (whether asymptomatically or symptomatically) (4). Sero-epidemiological surveillance can provide valuable information to policymakers for estimating serological prevalence in different age and sex groups and monitoring changes in prevalence and incidence in different geographical areas through conducting repetitive population-based surveys (5, 6).

In this way, the results of the interventions performed can be evaluated and new cost-effective interventions can be designed. The findings could be useful in providing comprehensive information on the history of infection and herd immunity and estimating the probability and timing of future epidemic waves (7). In addition, using the results of this type of study, it is possible to accurately estimate the infection fatality rate of the disease. In the absence of these studies, however, the estimated infection fatality rate of COVID-19 should be calculated based on the number of confirmed cases of the disease multiplied by a parameter that indicates the number of indeterminate or asymptomatic cases (8-12).

So far, all available evidence suggests that the consequences of the COVID-19 pandemic will not be limited to a short period of time and that the health system will need to plan for COVID-19 management in the months and possibly years ahead. Proper policymaking on effective interventions, prevention, and control of the epidemic requires accurate information on the burden and distribution of the disease across the country. Accordingly, the Ministry of Health and Medical Education (MOHME) of the Islamic Republic of Iran has designed regular nationwide population-based surveys of the Iranian COVID-19 Serological Surveillance (ICS) program, whose first survey was implemented in August 2020.

Aims

The objectives of the project are (i) to estimate the prevalence of COVID-19 in Iran at specific intervals, in total and by province, urban/rural area of residence, sex, and age groups; (ii) to estimate the monthly incidence rate of COVID-19 in Iran indirectly, in total and by province, urban/rural area of residence, sex, and age groups; and (iii) to determine the trend of changes in the prevalence and incidence of COVID-19 over time in the country and separately for each province.

Methods

Study Design

In the ICS program, nationwide population-based serological surveys for COVID-19 will be conducted regularly. In each survey, the target population will be all Iranians aged 6 years and over living in the country. Besides, non-Iranian citizens living in the country are also considered as a part of the target population.

Inclusion and Exclusion Criteria

The inclusion criteria of each survey will be as follows: (i) People with Iranian unique national identification number registered in the primary health care (PHC) electronic health record systems (SIB, SINA, and NAB); (ii) non-Iranian citizens without Iranian unique national identification number registered in the PHC electronic health record systems; (iii) aged 6 years and older; and (iv) sufficient physical ability to attend the blood sampling center.

Individuals hospitalized or isolated due to definite or probable diagnosis of COVID-19 who are in the contagious period will also be considered as the target population. If possible, blood samples will be taken from these people in full compliance with the health protocols at their place of residence. Otherwise, the reason for not including them in the study will be recorded, but they will also be taken into account in calculating the prevalence of the disease.

The exclusion criteria will be contraindications for venous blood sampling and unwillingness for participation in the study.

Sampling Design

To ensure adequate representativeness of the enrolled participants, stratified random sampling will be used for each survey, with each province being considered as a stratum. The Iranian and non-Iranian populations aged 6 years or older who are registered at the PHC electronic health record systems will be used as the sampling frame of the repeated surveys. In each province, sampling will be done through simple random sampling. The sampling process will be done independently in each survey.

Sample Size

The sample size of the provinces in the first survey will be equal. For Tehran province, the sample size will be as triple as other provinces to capture the large and diverse population living in Tehran province. To do so, each of the medical universities covering the population in Tehran province (Tehran, Shahid Beheshti, and Iran Universities of Medical Sciences) will enroll the sample size equal to other provinces. Non-Iranian citizens registered in the PHC electronic health record systems will be considered as a separate stratum.

For the first survey, the sample size of each province, 3 medical universities of Tehran province, and the population of non-Iranian citizens without a Iranian unique national identification number were calculated as 780, based on the results of the sero-prevalence of COVID-19 in Guilan province (33%) (13), relative estimation error equals to 10%, and a confidence interval (CI) of 95%. Considering a non-response rate of 10%, the final sample size in each province, each of the medical universities of Tehran province, and non-Iranian citizens' strata were considered 858; thus, a total of 29,172 individuals will be entered into the first survey. The sample size of other surveys will be calculated based on the results of the first survey; however, it will not be the same for different provinces.
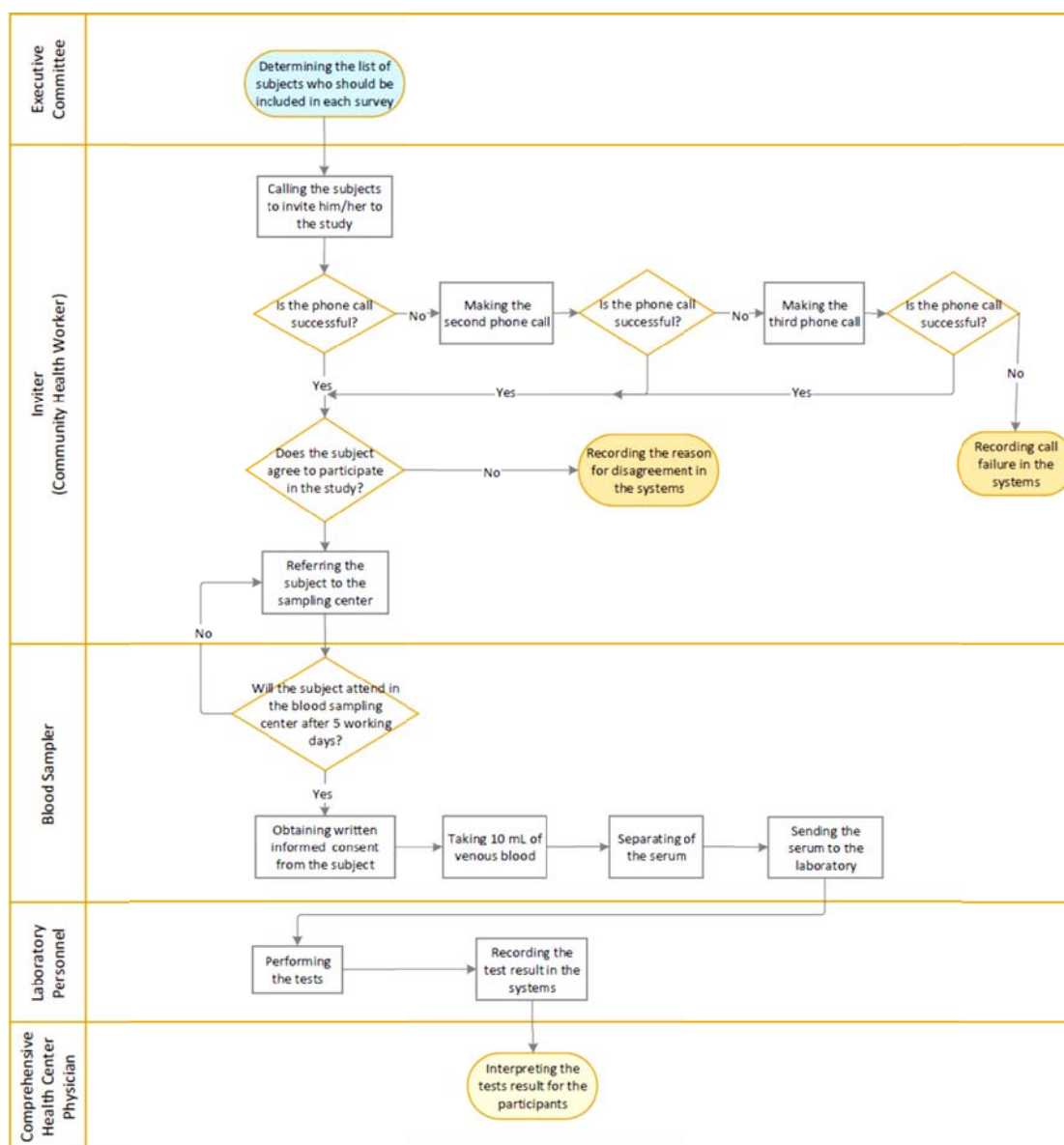

Fig. 1. Implementation Process Flowchart of the Survey

Inviting the Participants to the Districts' Blood Sampling Centers

Figure 1 shows the implementation process for each survey. After determining the list of enrolled people in the study for each university/faculty of medical sciences, their profile will be visible in the urban and rural community health workers interface in the PHC electronic health record systems. Community health workers, in coordination with the primary health care network of each district, will call the selected people and invite them to the selected blood sampling centers of the district for blood sampling. If they fail to make a phone call to the selected person, they keep calling up to 3 times and up to twice a day. In case of no response, lack of eligibility criteria, or dissatisfaction of the person for blood sampling, et cetera, such people are classified in the "non-responding" category and the reason for their non-response will be recorded in the PHC electronic health record systems by the community health workers.

Community health workers will provide the ground for maximum participation in the study by explaining the objectives of the study to participants. Individuals who give verbal consent to participate in the project are asked to refer to the selected blood sampling center of the district within a maximum of 5 working days with an identification (ID) card in hand and without any special preparation. If the participant does not refer to the sampling center after 5 working days, the community health workers will be informed through the systems and they will contact the participant again to refer. If the person does not refer despite the follow-up, the non-referral reason for such a person will be registered in the system.

Blood Sampling from the Participants, Serum Separation, and Transferring the Samples to the Laboratories

The blood sampling centers in each district will be selected so that the participants face with minimum risk of contracting the coronavirus and that they have easy access. Blood samples will be taken from the invited people in these centers. Blood sampling will be performed after obtaining written informed consent from the participants and in full compliance with health protocols. The sampler and the participant will both use personal protective equipment (mask, goggles, gloves, and appropriate gown). The sampling sites will be disinfected at intervals between visits with suitable disinfectants. A volume of 10 mL of intravenous blood will be taken from each person. Also, the unique code of the person in the PHC electronic health record systems, the name of the district and province, and the date of sampling will be labeled on the sample tubes. Up to 2 hours after blood sampling, the sample tubes will be centrifuged at 1000 to 1200 rpm for a maximum of 15 minutes, and then their serum will be transferred to plastic-sealed microtubes with an identification label and they will be stored at 4 C° to 8 C° in the refrigerator until the transfer time. Serum samples will then be transferred to the selected laboratory of the medical university in a 3-layer package at a temperature of 4 C° to 8 C° up to 24 hours after sampling.

Serological Testing Using ELISA Method on Serum Samples

After receiving serum samples from the blood sampling centers, the serological tests will be done on the samples by enzyme-linked immunoassay (ELISA) method using ELISA Kit SARS-CoV-2/IgG-IgM (Pishtaz Teb) according to the relevant protocol (Appendix 1) to determine IgM and IgG antibodies against COVD-19 in the selected laboratories of medical universities.

Recording Test Results in the PHC Electronic Health Record Systems

Once the results of the serology tests are known, the results will be recorded by the laboratory staff in the PHC electronic health record systems. The participants will be informed orally about the test results by community health workers. For proper interpretation of the results of the tests, all participants will be referred to a comprehensive health center physician. Information of people with a positive test can also be used in managing the epidemic. Print of the test results will be provided to the participants, if requested.

Providing Other Required Variables

In addition to the result of the tests, the following variables will be extracted from the participants' profile in the PHC electronic health record systems and added to the data of each survey: age, gender, province of residence, district of residence, name of comprehensive health center covering the individual, urban/rural area of residence, number of family members, having Iranian or non-Iranian citizenship, and history of positive COVID-19 PCR test.

Displaying the Results of Each Survey in the PHC Electronic Health Record Systems

At the end of each survey, the results, including estimation of the prevalence and monthly incidence rate of COVID-19 at the country and the province levels, will be displayed in the PHC electronic health record systems—in general, by urban/rural area, gender, and age groups. Also, the spatial distribution of the incidence and prevalence in the country and the trend of their changes over time in the provinces and the country will be displayed on the systems.

Estimating Sensitivity and Specificity of the Laboratory Kit

In this project, the Food and Drug Administration of Islamic Republic of Iran approved SARS-CoV-2 ELISA kits (Pishtaz Teb [catalogue numbers: PT-SARS-COV-2.IgM-96 and PT-SARS-COV-2.IgG-96]) to be used to measure IgG and IgM antibodies against COVID-19. Despite the report of the diagnostic accuracy of the kit by the manufacturer, as the accuracy of the kit is important in determining the prevalence of COVID-19, its accuracy will be reevaluated in a separate study.

To assess the sensitivity of the kit, 255 patients with a confirmed diagnosis of COVID-19 by the molecular method (PCR) will be tested (14). These individuals will

be a combination of those admitted to hospitals, outpatient clinics, and asymptomatic patients (15). Symptomatic patients will be selected among those whose symptoms had been appeared at least 3 weeks ago.

To estimate the specificity of the kit, 410 people with no COVID-19 disease will be included in the study. Serum samples of the biobank of Tehran Lipid and Glucose Study (TLGS) from 1 year before the onset of the COVID-19 pandemic will be used for this purpose (fall 2018 to summer 2019). The serum samples will be a combination of samples taken in all seasons of the year in equal proportions.

Education

In a virtual training session, the objectives and protocol of the study will be taught to the focal points of the health laboratories of the medical universities and they will be asked to transfer the training materials hierarchically by holding at least 1 virtual training session for community health workers, blood samplers, and technical laboratory personnel. Also, the study protocol, its attached guidelines, and instructions for implementing the project will be provided to all study facilitators.

Monitoring and Supervision

During each survey, all the study processes, including inviting randomly selected individuals to participate in the study, obtaining written informed consent, taking blood samples, separating serum from the samples, transferring samples from the blood sampling centers to the laboratory, performing serological tests, recording and interpreting the test results, and delivering the test results to the participants will be monitored. Two methods will be used for monitoring: (i) monitoring the progress of the study through the dashboard of PHC electronic health record systems based on the indicators in Table 1, and (ii) monitoring by the checklists. The directors of health laboratories of the medical university are responsible for the supervision of the optimal implementation of the study process. The monitoring checklists will be completed under their supervision.

Ethical Considerations

This project has been approved by the ethics committee of the National Institute of Health Research (NIHR) of the Islamic Republic of Iran (Ethics code: IR.TUMS.NIHR.REC.1399.019). The only intervention that will be performed in this project is taking 10 mL of venous blood for serological tests. Written informed consent will be obtained from all participants in the study. For participants aged 12 to 18 years, in addition to the individual, the consent of the parents or legal guardians will also be taken. In children under 12 years, written informed consent is obtained only from the parents or legal guardian. Blood samples will be anonymous and the test results will be communicated to participants through community health workers. Participants' information remains completely confidential and only their aggregated information

Table 1. Indicators for Monitoring the Progress of the Study and Their Calculation Methods

| Indicator | Definition | Proportions' Calculation Method | | Report Level |
		Numerator	Denominator	
Invitation performance	The number of people who have been successfully invited by a community health worker for blood sampling, and their percentage of the total number of people who should have been invited (determined sample size).	The number of people community health worker has successfully contacted and invited them to the study.	The total determined sample size in each level of the report.	Country, separately by province, medical university and district primary health care network, and in the subgroup of non-Iranian citizens
Blood sampling performance	The number of people from whom blood samples were taken, and their percentage of the number of people who were invited.	The number of people from whom blood samples were taken.	The number of people invited.	Country, separately by province, medical university and district primary health care network, and in the subgroup of non-Iranian citizens
Laboratory performance	The number of people whose serological test was performed and the result recorded, and their percentage of the number of people from whom blood samples have been taken.	The number of people whose serological test was performed and the result recorded.	The number of people from whom blood samples were taken.	Country, separately by province, medical university and district primary health care network, and in the subgroup of non-Iranian citizens
Total performance	The number of people whose serological test was performed and the result recorded, and their percentage of the total number of people who had to participate in the study (determined sample size).	The number of people whose serological test was performed and the result recorded.	The total determined sample size in each level of the report	Country, separately by province, medical university and district primary health care network, and in the subgroup of non-Iranian citizens

will be disseminated.

Data Management

As data collection and entering will be done using the Web-based application integrated into PHC electronic health record systems, many common sources of data entering errors will be prevented. Followed by the data collection stage of each survey, laboratory results along with other related information of participants in the PHC electronic health record systems will be used in the calculation of the prevalence and the incidence. All study participants have unique identifier codes linked to their unique national ID numbers but differ from them, so they remain anonymous. These codes will be used to exchange and communicate data.

The data screening steps will be performed by the analysis team comprises sorting, categorization, and checking the distribution of data. After finalizing the definitions and coding of the variables, all variables will be rechecked. In all stages of data management, no files will be changed and no variables will be deleted. Instead, the files will be numbered with changed or new variables.

Statistical Analysis

Data cleaning and statistical analysis will be conducted at the end of each survey. Prevalence of COVID-19 will be estimated at the national and provincial level by urban/rural area of residence, gender, and age categories (6-17; 18 to 39; 40 to 59; and ≥60 years). In the presence of either anti-SARS-CoV-2 IgG or anti-SARS-CoV-2 IgM, or both, a participant will be considered "positive." After correcting the false negatives and false positives of the serological test results, and weighting data based on the sampling design and response rate, a minimum biased estimate of the prevalence of COVID-19 will be obtained. Microsoft Excel (Microsoft Inc), STATA (Stata Corp), and R software (R Software Inc) will be used in statistical analysis. Prevalence proportions will be estimated with a 95% uncertainty interval (UI).

Statistical analysis of each survey will have the following steps: (i) correcting the crude prevalence proportions for the measurement error of the laboratory kit based on the sensitivity and specificity of the kit; (ii) converting adjusted prevalences of the first step to the individual data; and (iii) weighing the individual data of the previous step based on the sampling design and response rate. All of the above steps will be performed in 16 strata made up of a combination of 4 age, 2 gender, and 2 urban/rural categories in each province separately. The details of each of the above steps are described below:

1. Correcting the Crude Prevalence Proportions for the Measurement Error of the Laboratory Kit Based on the Sensitivity and Specificity of the Kit:

For this purpose, the classic probabilistic bias modeling or the Bayesian methods will be used.

(A) Correcting the laboratory kit's measurement error with the classic probabilistic bias modeling method:

One of the methods used for adjusting for the measurement error of the laboratory kit is probabilistic bias modeling (16-18). In this method, the beta probability distribution will be used to form 2 probability distributions, one for the sensitivity and the other for the specificity of the kit. Based on the sensitivity and specificity estimates and their standard errors obtained from the laboratory kit's performance study, the beta distributions of sensitivity and specificity will be constructed in such a way that their mean and standard deviation be matched with the point estimates and standard errors of the sensitivity and specificity obtained from the kit's performance study, respectively. For this purpose, for the beta distribution of sensitivity, the parameter α of the beta distribution is equal to the number of positive test subjects in the group of patients plus 1, and the parameter β of the beta distribution is equal to the number of people with negative test results in the group of patients plus 1. For the beta distribution of the specificity, the parameter α of the beta distribution will be equal to the number of people with negative test results in the healthy group plus 1, and the parameter β of the beta distribution will be equal to the number of people with positive test results in the healthy group plus 1. Then, the values of the α and β parameters of the constructed distributions of sensitivity and specificity will reach a value with a slight change in such a way that the mean and standard deviation of the constructed distributions will be the same as that obtained in the kit's performance study. Using the Monte Carlo sampling technique, 100,000 random samples will be drawn from the prior beta distributions of the sensitivity and specificity. In each sampling step, the adjusted prevalence will be calculated using equation (1) (9). The median, 2.5th, and 97.5th percentile of the posterior distribution of the adjusted prevalence will be computed as point estimate and 95% UI prevalence (18).

$$Adjusted.prevalence = \frac{crude.prevalence + spe - 1}{sen + spe - 1}$$

$$sen = p(Test+ | COVID+)$$

$$spe = p(Test- | COVID-) \tag{1}$$

One of the limitations of this method is that in 2 cases the adjusted prevalence estimates may be negative: (i) when the sum of the sensitivity and specificity of the test is <100%. In this case, the denominator of equation (1) becomes negative. This occurs when the accuracy of a test is so low that it is even worse than classifying patients by chance. (ii) When the specificity of the test is too low that the crude prevalence of the disease is less than the probability of false positives (specificity - 1). In this case, the numerator of equation (1) becomes negative. In such a case, it can be considered that all the positive cases detected by the test in that community are false positives. If we encounter the above limitation in the classic method, we will use the Bayesian method to estimate the prevalence proportions adjusted for the kit's measurement error.

(B) Correcting for the Laboratory Kit's Measurement Error With the Bayesian Method

In this method, the prevalence estimates adjusted for the sensitivity and specificity of the laboratory kit will not be negative (19). In the Bayesian method, in addition to the beta distributions of sensitivity and specificity, a uniform (0 and 1) distribution will also be considered for the crude prevalence proportion, and by the above three prior distributions, the posterior distribution of the prevalence, which is adjusted for the kit's measurement error, will be obtained (19):

$$prior_{prevalence} \sim U(0,1)$$
$$prior_{Sen} \sim scaled.Beta(\alpha, \beta)$$
$$prior_{Spe} \sim scaled.Beta(\alpha, \beta)$$

$$g(\theta|t,Sen,Spe) = \frac{\binom{n}{t}(c_1 + c_2\theta)^2 (1 - c_1 - c_2\theta)^{n-t}}{h(t)}$$

$$h(t) = \int_0^1 \binom{n}{t}(c_1 + c_2\theta)^2 (1 - c_1 - c_2\theta)^{n-t} d\theta$$

$$= c_2^{-1}\binom{n}{t}\{B(c_1 + c_2; t+1, n-t+1) - B(c_1, t+1, n-t+1)\} \quad (2)$$

In equation (2), Θ is the crude prevalence, t is the number of people with positive results from n people tested, $C_1 = 1 - Spe$, and $C_2 = Sen + Spe - 1$ (19).

2. Converting the Adjusted Prevalences of the First Step to the Individual Data

In the first step, the adjusted prevalence for the kit's measurement error will be calculated using the crude prevalence, sensitivity, and specificity of the tests in each of the 16 subgroups for each province separately. Because the prevalence is a summary measure, the data must be individualized before weighting. For this purpose, in each of the 16 subgroups of each province, using the number of participants and the adjusted prevalence of that subgroup, the data will be simulated individually using a binomial distribution.

3. Weighing the Individual Data of the Previous Step Based on the Sampling Design and Response Rate

(A) Weighting to Obtain the Provincial Prevalence Estimates

When calculating the provincial prevalence estimates, to correct the differences in the age-sex-urban/rural area distribution of the study sample, with their distribution in the population of each province and the effect of nonresponse, the following 2 weights will be used:

1. The weight of correcting the differences in the age-sex-urban/rural area distribution of the study sample, with their distribution in the population of the province: This weight is the inverse of the ratio of the number of samples determined for each age-sex-urban/rural categories by the population of that category in each province, based on the population projection for 2020 by the Statistics Center of

Iran.

2. The weight of responding: This weight will be used to correct the effect of nonresponse in the prevalence estimates. As the variables of age, sex, and urban/rural and the province of residence may be effective on participation in the study, the weight of responding will thus be obtained by dividing the number of determined samples by the number of the participants in each of the 16 age-sex-urban/rural categories (inverse of the probability of responding).

Multiplication of the 2 mentioned weights will be used to calculate the prevalence estimates in each province.

(B) Weighting to Obtain the National Prevalence Estimates

In order to calculate the final corrected prevalence proportions at the country level, it is necessary to consider the sampling design. Therefore, the sampling weight will be calculated by dividing the number of the population of each province by the sample size of that province. To calculate the national prevalence estimates, this weight will also be multiplied by the previous 2 weights.

Indirect Estimating of COVID-19 Incidence Rate

The monthly incidence rate of COVID-19 could be estimated from data of 2 consecutive Surveys (20). The approximate incidence rate (IR) based on the prevalence proportions of first and second surveys will be estimated with equation (3) (20); the 95 % UI of the monthly incidence rates will be computed by the bootstrap method:

$$IR = \frac{2*(pos_2 - (s_1 * pos_1))}{(N_1 - pos_1) + (N_2 - pos_2)} \quad (3)$$

s_1: the proportion of the positive individuals from the first survey who survived until the second survey (s = survivor/ number of positive individuals of the first survey);
pos: number of positive individuals at each survey; and
N: total number of participants at each survey.

Strengths and Limitations

The ICS program surveys are the largest population-based surveys in Iran that will be conducted nationally to estimate the minimum biased prevalence of COVID-19 infection. The use of serological tests allows the diagnosis of infection in asymptomatic people and also in those with mild symptoms. Therefore, by conducting these surveys, a clearer picture of the prevalence of COVID-19 infection will be obtained in the country. Repeating the surveys will allow policymakers to observe changes in the magnitude of the infection over time. In this project, the frequency of the infection in the population of non-Iranian citizens (as a high-risk population) will also be determined and monitored over time. We will also correct the measurement error of the laboratory kit used in the surveys.

One of the limitations of this project is the use of the registered population in PHC electronic health record systems as a sampling framework. Despite the systems coverage of over 90% of the population in many provinces of

the country, in some provinces such as Tehran, the coverage of the systems is about 80%. Fortunately, this coverage is being improved rapidly as the MOMHE implements other active surveillance programs for COVID-19. We will address this limitation to some extent through weighting during statistical analysis. Another limitation of our project is the effect of such factors as the severity of symptoms, interval from the onset of the infection, age, et cetera on the sensitivity and specificity of COVID-19 serological tests. When estimating the sensitivity and specificity of the tests, we will try to use a sample that includes people with all modes of variables that affect the validity of the tests; and when analyzing the results, we will use the probability distribution of sensitivity and specificity to correct the tests measurement error. Another limitation of this project could be the low participation of people who were invited to the study. Efforts will be made to minimize nonresponse by using the capacity of the PHC networks and the credibility of the health system among participants as well as regular follow-ups and staff training. The possible nonresponse bias will be adjusted using weighting based on the inverse probability of responding.

Conclusion

The implementation of these surveys will provide a comprehensive and clear picture of the magnitude of COVID-19 infection and its trend over time for health policymakers at the national and subnational levels; in addition, it provides a suitable ground for evaluating the interventions and designing new interventions at local levels.

Acknowledgment

Conflict of Interests

The authors declare that they have no competing interests.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020.
2. Theel ES, Slev P, Wheeler S, Couturier MR, Wong SJ, Kadkhoda K. The role of antibody testing for SARS-CoV-2: is there one? J Clin Microbiol. 2020.
3. Bryant JE, Azman AS, Ferrari MJ, Arnold BF, Boni MF, Boum Y, et al. Serology for SARS-CoV-2: Apprehensions, opportunities, and the path forward. Sci Immunol. 2020;5(47).
4. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet. 2020.
5. Bajema KL, Wiegand RE, Cuffe K, Patel SV, Iachan R, Lim T, et al. Estimated SARS-CoV-2 Seroprevalence in the US as of September 2020. JAMA Intern Med. 2020.
6. Hallal PC, Hartwig FP, Horta BL, Silveira MF, Struchiner CJ, Vidaletti LP, et al. SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. Lancet Glob Health. 2020;8(11):e1390-e8.
7. Bryant JE, Azman AS, Ferrari MJ, Arnold BF, Boni MF, Boum Y, et al. Serology for SARS-CoV-2: Apprehensions, opportunities, and the path forward. Sci Immunol. 2020;5(47):eabc6347.
8. Spychalski P, Błażyńska-Spychalska A, Kobiela J. Estimating case fatality rates of COVID-19. Lancet Infect Dis. 2020.
9. Bendavid E, Mulaney B, Sood N, Shah S, Ling E, Bromley-Dulfano R, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. MedRxiv. 2020.
10. Kimball A. Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility—King County, Washington, March 2020. Morb Mortal Wkly Rep. 2020;69.
11. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance. 2020;25(10):2000180.
12. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science. 2020;368(6490):489-93.
13. Shakiba M, Nazari SSH, Mehrabian F, Rezvani SM, Ghasempour Z, Heidarzadeh A. Seroprevalence of COVID-19 virus infection in Guilan province, Iran. medRxiv. 2020.
14. Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. Cochrane Database Syst Rev. 2020;6.
15. To KKW, Tsang OTY, Leung W-S, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis. 2020.
16. Khalagi K, Mansournia MA, Motevalian SA, Nourijelyani K, Rahimi-Movaghar A, Bakhtiyari M. An ad hoc method for dual adjusting for measurement errors and nonresponse bias for estimating prevalence in survey data: Application to Iranian mental health survey on any illicit drug use. Stat Methods Med Res. 2018;27(10):3062-76.
17. Khalagi K, Mansournia MA, Rahimi-Movaghar A, Nourijelyani K, Amin-Esmaeili M, Hajebi A, et al. Assessing measurement error in surveys using latent class analysis: application to self-reported illicit drug use in data from the Iranian Mental Health Survey. Epidemiol Health. 2016;38:e2016013.
18. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data: Springer Science & Business Media; 2011.
19. Diggle PJ. Estimating prevalence using an imperfect test. Epidemiol Res Int. 2011;2011.
20. Hallett TB, Zaba B, Todd J, Lopman B, Mwita W, Biraro S, et al. Estimating incidence from prevalence in generalised HIV epidemics: methods and validation. PLoS Med. 2008;5(4).

Appendix 1. Guide to Calculating IgG and IgM Results in the Laboratory

Guide to Calculating IgG Results:

Any ELISA reader device with the ability to measure light absorption at a wavelength of 450 nm can be used.

(1) Read the light absorption of the controls and samples with an ELISA reader device at 450 nm and, if possible, in front of a 630 nm reference filter.
(2) Reduce blank light absorption from controls and samples.
(3) Calculate the cutoff value according to the following formula:

$$\text{Cutoff value} = \text{Mean of negative control light absorptions} + 0.15$$

To determine the positive and negative answers, obtain the index value by dividing the sample light absorption by the cutoff value:

$$\text{Cutoff Index (COI)} = \text{OD of sample/cutoff value}$$

According to this formula, values >1 are considered positive, and values ≤1 are considered negative.

Guide to Calculating IgM Results

Any ELISA reader device with the ability to measure light absorption at a wavelength of 450 nm can be used.

(1) Read the light absorption of the controls and samples with an ELISA reader device at 450 nm and, if possible, in front of a 630 nm reference filter.
(2) Reduce blank light absorption from controls and samples.
(3) Calculate the cutoff value according to the following formula:

$$\text{Cut Off value} = \text{Mean of negative control light absorptions} + 0.25$$

To determine the positive and negative answers, obtain the index value by dividing the sample light absorption by the cutoff value:

$$\text{Cutoff Index (COI)} = \text{OD of sample/Cut-off value}$$

According to this formula, values >1 are considered positive, and values ≤1 are considered negative.

Warnings and Recommendations of the Manufacturer of the Kit

According to the results obtained from the implementation of the Periodic Maintenance Service (PMS) carried out by Pishtaz Teb, we hereby draw the opinion of respected laboratories to the following points:

❖ In the COVID-19 ELISA kits of this company, the antibodies are measured indirectly, which according to the instructions of the kit, it is necessary to dilute the serums before the test, in this regard, not enough attention to the preparing the samples according to the kit instructions can lead to false and especially false positive answers close to cutoff.
❖ Because of the high sensitivity of the antigen and antibody reaction in the company's COVID-19 kits, there is a possibility of a rapid reaction if positive samples enter the negative wells (carry over) and considering the possibility of this occurring during washing, particularly handwashing using a piset, syringe, or pipette, it is recommended that washing be done with caution to prevent this complication.
❖ Therefore, it is recommended to repeat the test more with more care if weak positives are observed close to the cutoff.