Supplementary material: Example of protocol application

In this practical example is shown how the proposed base protocol could be used to calculate survey sample sizes that could serve the double purpose of estimating the “true” prevalence (TP) of SARS-CoV-2 infected individuals (both symptomatic and not) and, if no cases are found, reaching the confidence in freedom (PFree) at the cut-off design prevalence Pu. The PFree is the confidence that, if no detection occurs, the disease is anyway below the hypothetical design prevalence Pu. N.B. The mentioned inputs are simulated, and must be set using realistic values, by the task force unit applying the protocol. The mentioned references and figures can be found in the article. The Ausvet links [15-16] are used just as an example.

Infection scenario:

One person has been found positive to SARS-CoV-2 by oral swab using a PCR test, which is assumed to have sensitivity Se = 95% and specificity Sp = 100%. Thus we could be between steps A and B in Fig. 1.

Information from anamnesis and tracing investigations:

From the anamnesis, the patient showed mild respiratory symptoms (some dry coughing) during the last 5 days. From the tracing investigations it is found that, the person came back two weeks ago, from an high risk area where a high number of COVID-19 cases have been reported during the last month. The tracing of local contacts within the (home) country allows preventive quarantine of all people that the suspected index case remembers to have met. Testing of quarantined people is also carried out, as risk based surveillance. Moreover, it is found that during the last two weeks the patient has visited clubs, gym, shops and working places in his commune, where approximately 5,000 people (Np) are resident. After some days, a total of 49 people result positive to the virus (traced contacts and others), and all of them are resident in the same area. All the apparently secondary cases lived and moved around in the same residential zone delimited by the task force, before the quarantine was enforced.

Research question from the authorities:

The virus could have been spreading before and even after the first patient was found:

I. If disease spreading already occurred in the area, what could be the current true prevalence of infected people (asymptomatic and not)? TP?

II. If no further cases are detected what is the confidence that the remaining target population (Np) is free from disease and that the 50 positives where the only ones infected (thus Pu < 1% in the remaining population)? PFree?

To calculate the sample size (n1) and to address question “I” we could use the Ausvet link [15] (Fig. 3) and set:

a) The (a priori) assumed true prevalence PriorPT = 1.0%. We could set this input in this way ,because this is the apparent prevalence (AP) we observed (50/5000 = 1%) before the survey is carried out. Alternatively, this input could be set according to experts or data from other infected areas.
b) The Se and Sp of the test are set = 95% and 100%, respectively.
c) The precision is set to 2%
d) The confidence level is set to 98%.

To calculate the sample size (n2) and to address question “II” we could use Ausvet link [16] (Fig. 3) and we set:

a) The test sensitivity Se = 95%
b) The prior confidence in freedom (PriorPFree), before the survey, to 50%
c) The probability of further disease introduction (PIntro) from outside the area to 0%
d) The aimed confidence in freedom (PFree) to 95%
e) The size of the population (Np) = 4950

f) The design prevalence of infected people Pu = 1.0% at which we aim detecting at least one positive with population sensitivity SSe, if such cut-off has been reached in reality. If the infected already detected in the tracings were the only cases (50), and if no others are detected through the survey, the PFree is the confidence that Pu is < 1% in the remaining Np population.

3. Outputs

In Table 1 are shown the sample sizes that would be needed to address research question "I" and to estimate the TP. Sample sizes are provided for different combinations of precision and assumed PriorTP, while keeping Se and Sp fixed to 95% and 100%, respectively. If assuming PriorTP = 1% and precision 2%, then n1 = 142 people should be tested randomly; while for precision 1% n = 565. The program produces also sample sizes that are higher than the overall population of residents (Np) because, adjustment for "finite" population size has been removed [15]. If the protocol is applied to "finite" populations, adjustments could be applied [14].

Table 1. Sample size required to estimate the true prevalence TP within the target population Np, using a test with Se = 95% and Sp = 100%, and considering varying levels of precision and prior true prevalence (PriorTP)

	PriorTP = 0.01	**PriorTP = 0.02**	**PriorTP = 0.05**	**PriorTP = 0.1**	**PriorTP = 0.2**	**PriorTP = 0.5**
Precision = 0.01	565	1118	2714	5156	9229	14954
Precision = 0.02	142	280	679	1289	2308	3739
Precision = 0.05	23	45	109	207	370	599
Precision = 0.1	6	12	28	52	93	150
Precision = 0.2	2	3	7	13	24	38

Regarding question "II", it seems that to achieve the aimed confidence in freedom (PFree = 95%) 298 people should be tested randomly in the target Np. By using this sample size the achieved confidence in

detecting at least one infected person (surveillance sensitivity SSe) would be 94.7%, if by the day of sampling, the within area design prevalence Pu = 1% has already been reached.

Overall interpretation of survey´s output according to samples sizes n1 and n2 used

If testing randomly 298/4950 persons and if at least one is positive, the percentage of infected would represent the TP. If no cases are detected, the confidence that the TP is below 1% would be around 95%, because if such a prevalence was reached in reality, the testing scheme of the survey had SSe = 95% chances to detect at least one positive and reject the hypothesis of “freedom”.

With the results of the first survey we can see in which step (of Fig. 1) the epidemic could be. Then eventual risk/control measures could be addressed (Fig. 2). Moreover, according to the stage of the outbreak it could be decided if further surveys are needed, and in that case, how many and how often.