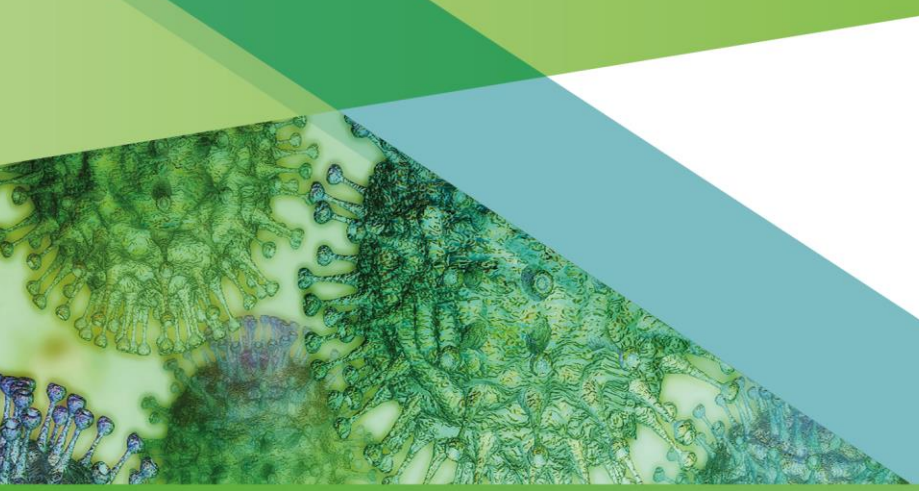

TECHNICAL REPORT

Methodology for estimating point prevalence of SARS-CoV-2 infection by pooled RT-PCR testing

28 May 2020

Scope

The scope of this document is to provide EU/EEA Member States and the UK with technical guidance for estimating the prevalence of SARS-CoV-2 infected cases, through a cross sectional study design based on pooled sampling for RT-PCR testing on a random population sample.

The proposed methodology is for surveillance purposes and does not address pooled testing for clinical (i.e. diagnostic) or public health (e.g. contact tracing, infection prevention and control, screening) management of individuals.

The goal is to provide a tool to monitor the activity of the disease with high efficiency in terms of laboratory resources. The methodology provides a way to estimate prevalence from the results of a pooled RT-PCR test, without the need to identify individual test results, and the statistical foundation for the strategy. We further provide a method to optimise the pooling strategy with input for sample size decisions. Finally, we also provide an in-house R package to help public health authorities define a strategy and analyse the results.

Target audience

Public health authorities in the EU/EEA Member States and the UK. The set-up of a PCR prevalence study requires close collaboration between epidemiologists and virologists.

Background

The COVID-19 pandemic is putting a considerable strain on societies due to the large morbidity and mortality, the profound impact on healthcare systems, and the economic and societal harm associated with the physical distancing measures [1]. To better control the disease, allocate healthcare resources and evaluate the effectiveness of control measures, it is important to achieve accurate estimates of the prevalence of infection at the national and sub-national levels. Current surveillance strategies, such as mandatory notification and sentinel surveillance by general practitioners or in hospitals, are mainly based on the testing and reporting of symptomatic individuals or COVID-19 related deaths, and therefore underestimate the spread of the virus in the pauci- and asymptomatic population or in specific risk groups with limited access to healthcare or testing.

Suggested citation: European Centre for Disease Prevention and Control. Methodology for estimating point prevalence of SARS-CoV-2 infection by pooled RT-PCR testing. Stockholm: ECDC; 2020.

Furthermore, sentinel systems are less efficient in countries recommending patients not to visit general practitioners if they have COVID-19 related symptoms, countries where stay-at-home policies are in place, or countries where alternative testing sites and policies have been introduced.

For all these reasons, estimation of point prevalence of acute infection through PCR testing in a cross sectional survey design on a random population sample, could provide valuable information about the pandemic situation. Such studies could be repeated over time and be complementary to the cumulative prevalence estimations from sero-epidemiological studies or to other surveillance systems. The point prevalence found can further be used to feed into SEIR models of the outbreak, to improve estimates of the epidemic evolution and to assess the effect of control measures. In the context of large testing needs and upscaling test capacity, it is also fundamental to develop study protocols that improve efficient use of testing resources.

A few Member States have performed point prevalence studies in the general population:

Vo, Italy, February 2020 [2]: Two surveys were conducted, one to two weeks apart, in an Italian town with a population of 3 275 individuals. In the first survey, which was conducted around the start of the lockdown of the town, an infection prevalence of 2.6% (73/2 812) was found (95% confidence interval (CI) 2.1–3.3%). In the second survey, which was conducted at the end of the lockdown, a prevalence of 1.2% (29/2 343) was found (95% CI 0.8–1.8%).

Iceland, March 2020 [3]: Two surveys were conducted in Iceland: one using an open-invitation method (convenience sample) to the entire population of Iceland, and one using a random population sample. Of those tested in the open-invitation screening, 0.8% tested positive. For the random population screening, 6 782 randomly chosen Icelanders between 20 and 70 years of age were invited to participate through a telephone text message sent between March 31 and April 1. Of these invitees, 2 283 (33.7%) had participated by 4 April. Thirteen (0.6%) tested positive for the virus.

Sweden, March – April 2020 [4]: The Public Health Agency of Sweden, in collaboration with Swedish Armed Forces, conducted a prevalence survey in **Stockholm**, Sweden. Between 26 March and 3 April 2020, 1106 persons (adults or parents of selected children from a predefined representative study panel) were invited, of whom 738 participated (age range from 2 to 86 years of age). No analysis could be done for 31 of the participants. Eighteen out of 707 analysed samples, or 2.5 % (95% CI 1.4–4.2%), tested positive with PCR for SARS-CoV-2 infection. Under the assumption that this was a non-biased sample, this result would suggest 59 364 infected persons in Stockholm in the last week of March. The study protocol was repeated with 679 participants between 21 and 24 April and resulted in an estimate of 2.3% [5]. A **nation-wide** study revealed that based on 2 573 persons tested, 0.9 % of the population was infected between 21 and 24 April [5].

United Kingdom, April – May 2020 [6]: A partnership including the Office for National Statistics, the University of Oxford, the University of Manchester, Public Health England, and the Wellcome Trust conducted a COVID-19 infection survey. In the pilot survey, from 27 April to 10 May, based on 33 individuals testing positive out of 10 705 people, an average of 0.27% (95% CI 0.17-0.41%) of the community population was estimated to be infected. Acknowledging the unknown false positive and false negativity rate of the tests, the results are reported as 'positivity rate' rather than as 'prevalence rate'.

All these studies were based on individual testing of samples, which makes them expensive to reproduce on large scales.

We propose a guidance for efficiently conduct large-scale population PCR-RT based prevalence studies through pooled group testing. The technique is based on the grouping of clinical specimens of different people in diagnostic pools before testing. Only a single test is then performed on each pool. Assuming perfect sensitivity and specificity, a positive result indicates one or more affected individuals in the pool and a negative result indicates no positive individuals. While pooling can apparently introduce bias in prevalence estimates, it's accuracy becomes comparable to individual testing when disease frequency is low (< 33%) [7,8].

One of the first applications of pooled group testing was in the screening for syphilis the US army in the 1940s, but then it became a standard technique in the laboratory diagnostic context [9], the quality control field [10], and the control of the spread of infectious diseases, e.g. malaria [11,12]. Lately, the technique has been studied in the context of SARS-CoV-2 infection due to the need to perform a large number of RT-PCR tests to identify cases [13-16].

For the sole purpose of a prevalence estimation study, no individual test results need to be obtained. However, to obtain individual results for the implementation of contact tracing or infection prevention and control measures, subsequent confirmatory tests can be done. For this, the original samples must be stored until the pooled test is performed and smaller sample pools should be considered according to the laboratory method validation. It is important to note that our pooling method is not recommended for diagnostic purposes. Guidance on the correct implementation of individual cases identification through a group testing approach is outside the scope of this document.

Study design and implementation

The testing strategy proposed here is general and can be applied in a number of epidemiological situations. We will describe some possible applications related to the current COVID-19 epidemic.

Study populations and sampling strategies

Study organisers need to identify the study population and develop a sampling strategy aimed at acquiring a sample which is as unbiased as possible. Subgroup analysis is not possible with the proposed analytical strategy. If random population samples are already drawn for other studies, e.g. sero-epidemiological studies, swabbing for PCR testing could be added, provided the sample sizes are large enough. Country expertise should be utilised to select the best strategy.

The goal of this sampling strategy is to obtain a representative sample of the population. For this purpose, it is important that the selection of and willingness to participate of individuals is unrelated to their health status (symptomatology) and to that of their contacts and exposures. For example, hospital-based or general practitioner based sampling would produce an upward bias in prevalence, while recruiting participants at a fixed public location (e.g. entrance of shops), pregnant women, blood donors, etc. would create a downward bias by selecting healthier parts of the population.

We describe below two general sampling strategies; geographical sampling and risk group sampling.

Geographical sampling

The focus is the assessment of the prevalence of COVID-19 in a geographical area of interest, which could be a region or a city. Larger areas may be associated with too high heterogeneity for a prevalence estimate to inform public health action or serve as a surveillance indicator. For countrywide prevalence monitoring, study organisers may stratify by regions with each sufficient sample size, or perform multiple region-specific studies.

Ideally, a (systematic) random sampling is taken from a population register. If available, predefined representative study panels (set up for other or general study purposes) could be used.

If no population register or study panels that represent the general population are available, investigators may select individuals from randomly selected households. In this case, we suggest the following procedure to avoid bias in choosing an individual from a household:

- use systematic random sampling to select a number of households equal to the number of individuals needed;
- acquire specimens from all individuals of the household;
- sub-sample randomly individual material, independently from the household membership, until reaching the needed number of samples.
- pool the material, independently from the household membership, and perform RT-PCR testing on pools.

With this procedure, residual clustering effect due to household membership is expected to be negligible due to the limited household size compared to the full sample size. Clustering effect may become relevant when including large household structures like residential home, hospitals and prisons, especially if the final sample is comparatively small. It is advisable not to include these groups in the study, but instead perform separated risk group prevalence studies (cfr. Risk group sampling).

Household based sampling will be particularly effective when stay-at-home policies are in place as it is more likely that all members of the household will be at home, reducing the risk of introducing bias in the selection.

If both registry based and household based sampling prove logistically unfeasible, other strategies based on convenience sampling, like an open invitation recruitment, e.g. via a website or social media, or recruitment at public places, may be used, at the cost of bias.

Independent of the sampling strategy, information (e.g. age, sex, number of household members, occupation, socio-economic status indicators, presence of symptoms) may be registered for analysis of representativeness of the population.

Risk group sampling

The proposed methodology can be utilised to assess the prevalence of infection in specific populations. Risk groups of interest could be healthcare workers, the elderly, hospitalised patients, etc. Also in this design, it is important to strive for selection of individuals independent of the perceived disease status. The objective of this protocol is to estimate population prevalence and participants need to be properly informed.

The proposed methodology should not be used to replace guidance specific to clinical or public health management of individuals or institutions of risk groups (e.g. comprehensive testing of residents and staff in long term care facilities), but may be used to complement other forms of surveillance [17].

The methodology does not preclude public health action based on the findings, e.g. in terms of infection prevention and control, contact tracing, or reinforcement of testing policy and implementation.

If a clustered sampling design is used, e.g. selecting individuals from a limited number of institutions, the clustering effect may be relevant. In this case, it may be more appropriate to perform separate sub-studies for each institution and then combine the results.

Study frequency

Such prevalence studies can be performed on an ad hoc basis or they can be repeated at regular intervals to provide surveillance of SARS-CoV-2 prevalence. The periodicity of studies should consider the public health objectives and available resources. Pooled analysis can greatly reduce the costs in laboratory material, but does not affect the logistic, human and time costs of the study.

For surveillance of prevalence, we suggest performing studies with enough periodicity to closely follow the evolution of the epidemic. Given the rapid spread of SARS-CoV-2, a window of two to three weeks would be ideal. The utilisation of pooled testing makes it possible to reduce the number of RT-PCR tests performed, while the Bayesian inference methods (cfr. Analysis methods) make it possible to utilise the results of precedent studies to decrease the required sample size.

For risk group analysis, if the goal is estimating a difference in risk and not surveillance, frequent repetition of studies is less important, so the organisers should focus on collecting an adequate number of samples and avoid biases.

Sample size

As described later, we provide tools aimed at the optimisation of the pooled test design (i.e. choice of the pool size and number of pools) which also provide information on the impact of the total sample size on the study results.

The sample size needs to be sufficiently large to address non-response. From the experience of the above mentioned studies, we suggest considering a non-response rate of about 30%, but the actual rate will largely depend on the practical set-up of the study. In addition, non-response could be higher in specific sub-groups, and based on previous experience, some specific sub-groups may be over-sampled [18].

Participant invitation

The invitation could be sent via a postal invitation, e-mail invitation, phone call, or SMS, according to what is available in the registries.

Organisers need to consider the balance between the efficiency gains of pooling against the risk of non-response due to not being provided with an individual result. If no individual test results will be obtained, this needs to be clearly specified in the informed consent statement (as well as in the study protocol and data and material agreement). Participants may still be provided the study results as a motivation. Furthermore, participants may be provided with a test opportunity outside the study (e.g. via a voucher, or via a dedicated telephone number).

Specimen collection

For specimens taken by healthcare personnel, we propose the use of nasopharyngeal and/or oropharyngeal swabs. If both are taken, they can be mixed together in one specimen [19].

Alternatively, participants may perform self-swabbing of throat and nose (nostrils). Also saliva can be collected but requires a separate specimen and the use of this alternative sample type is still under evaluation [20-22].

We propose three possible methodologies for specimen's collection.

- **Self-swabbing:** Once people have accepted to participate, self-swabbing kits can be delivered with appropriate instructions on self-swabbing. The Public Health Agency of Sweden recommends this methodology in their report [4]. Video tutorials are an effective way to perform a successful training and can be added to written instructions. Instructions need to be given on how, when and where the specimens from the individual (or from all members of the household) will be collected. A safe pick-up and transport can be organised by an appropriate courier or via a third party collaboration, e.g. the CBRN Defence Unit in the Swedish study. Alternatively, people may be asked to bring the specimen (with instructions) to a given location, e.g. triage or testing point.
- **Collection of specimens at home:** Trained personnel, with due personal protective material, visit selected individuals/household to perform the swabbing and collect the material. The personnel do not necessarily have to be healthcare workers, but need to be appropriately trained, similar to regular testing sites.

- **Collection of specimens at established or ad hoc testing centres:** Alternatively, the selected individuals/households may be asked to present at testing sites with appropriate instructions and coordination. This solution may increase the risk of non-response bias.

Testing procedure

For prevalence studies, the pooling of clinical specimens may be preferred over the pooling of RNA, because of possible shortages of RNA extraction kits that have been described by EU/EEA countries [23]. RNA extraction can thus be done once for each pool of clinical specimens.

RT-PCR is a very sensitive technique and depending on the Limit of Detection (LoD) of the individual assay, it should be able to detect virus in pooled samples with low viral load [23]. However, increasing the number of pooled specimens will increase the probability of false negative results especially in samples with low viral load.

Storage and mixing of clinical samples can be done in any format. As an example, 96 2.5ml deep-well plates, with 12-well columns or 8-well rows can be used to mix the individual samples. Multi-channel pipettes enable simultaneous multi-sample pipetting manually. The final volume should be mixed thoroughly. Further pooling of the initial pools can proceed. The final arrangement will require the distribution of the individual specimens in pools in equal numbers (e.g. eight individuals per pool according to the laboratory's assay validation) with each individual material being present in not more than one pool.

RNA extraction and RT-PCR

After clinical sample pooling, RNA extraction can follow, using standard volumes based on the RNA extraction protocol. RT-PCR can follow according to the laboratory's validated protocols. Extension of the RT-PCR cycles and modification of the protocol (e.g. to decrease the amount of buffer) may be considered to improve the detection limit and enable increasing the number of pooled samples, but this should be validated in the individual laboratory [24]. The limit of detection largely depends on the individual protocols and therefore any modifications will need to be validated in the laboratory.

Many SARS-CoV-2 testing laboratories are using robotic pipetting platforms to rapidly and accurately pipet for RNA extraction or RT-PCR assays in a high throughput manner. These robotic systems may be adjusted to perform pooling of clinical specimens. However, we suggest that the pooling strategy is validated using manual pipetting before using robotic platforms.

Assay validation

Validation of the level of pooling will need to be done in the individual laboratory based on the RNA extraction and RT-PCR methodology used. Validation can be done by mixing the intended number of known positive and negative clinical specimens and, following procedures, determine the maximum number of specimen pooling that does not affect the method's accuracy and enables detection of all of the positive samples in the pool [25]. The maximum number of pooled samples (to obtain >95% sensitivity) depends on the laboratory's protocols and especially the Limit of Detection of the RT-PCR assay. It has been suggested that the pooling doesn't affect PCR sensitivity for SARS-CoV-2, with theoretical and experimental studies suggesting a high (> 95%) sensitivity up to 32 samples per pool [7,24,25], but this will largely depend on the above mentioned validation.

Assay results

Group testing will give the following results:

- a negative result will indicate that none of the individuals of the pool is infected;
- a positive result will indicate that at least one individual in the pool is infected.

Biosafety precautions

Manual mixing may increase risk of operator exposure to SARS-CoV-2. Therefore, strict biosafety precautions are needed. Please refer to the ECDC webpage for more information on biosafety precautions [19].

Statistical analysis methods

Estimation of prevalence from pooled test results

The estimation of population prevalence from pooled test results can be achieved in a number of ways [26-28]. We present three methodologies that allow incremental refinement in the estimations but with increasing technical requirements. To overcome implementation difficulties, we provide an R package (https://github.com/EU-ECDC/PooledPrevalence) with tools to obtain state-of-the-art prevalence estimates and guide study design. Similar packages for pool testing-based prevalence estimation are also available from other authors [29]; our package offers the addition of a formal Bayesian solution to the problem.

The definitions use following notation: p estimated COVID-19 prevalence; p_{test} probability of a PCR test on a pool of being positive; w number of pools tested in a study; k number of positive pools in a study; s pool size (number of individual samples in a pool).

Maximum likelihood estimation

This methodology is the most simple to implement and easy to reproduce with most statistical programs. It cannot be used in extreme cases in which there are no positive pools for a given study, which is something that may arise in the case of underpowered studies; in that situation assuming $k = 0.5$ will allow to have results at the cost of an upward bias in prevalence.

Given a test result $p_{test} = k/w$, the most likely prevalence is estimated as [30]:

$$(1)\ p = 1 - \sqrt[s]{1 - k/w}$$

and the relative 95% Confidence Intervals (95CI) as:

$$(2)\ CI_{95\%} = 1 - \sqrt[s]{1 - logit^{-1}(logit(p_{test}) \pm 1.96\ \frac{1}{\sqrt{wp_{test}(1-p_{test})}})}$$

with $logit(x) = \log(\frac{x}{1-x})$, and the logistic function its inverse $logit^{-1}(x) = \exp(\frac{\exp(x)}{1+\exp(x)})$.

Bayesian conjugate estimation

This methodology allows the possibility to have robust estimates in the presence of zero positive pools in a study. The model produces the posterior probability distribution of the prevalence given a test result. Since the probability of observing k positive pools follows a binomial distribution, it is possible to infer the posterior probability of the observed results p_{test} employing the posterior conjugate distribution, which is a Beta with shape parameters α, β (prior distribution cfr. later) incremented respectively with the number of positive pools k and negative pools $w - k$. Then (1) is used to estimate the prevalence from the obtained probability distribution. Since the whole posterior distribution is estimated, it's possible to extract the desired probability quantiles q. The usual quantiles are the median ($q = 50\%$) and the 95% Credibility Interval (95%CrI) given by $q = 2.5\%$ and $q = 97.5\%$:

$$(3)\ p_q = 1 - \sqrt[s]{1 - qBeta(q,\ \alpha + k, \beta + w - k)}$$

with $qBeta(q, x, y)$ being a function that returns the q quantile from the posterior conjugate Beta distribution with parameters $x = \ \alpha + k$ and $y = \ \beta + w - k$.

Bayesian hierarchical estimation

This method can specify all the uncertainties by building a generative model of the observed results. It can take into account the loss in sensitivity related to the swabbing (probability that viral material is obtained during the swabbing of infected participants). The model can be also expanded to take into account the possible loss in sensitivity due to pooling and the over-dispersion that may derive from the existence of independent risk groups in the population. The current version of the algorithm assumes perfect PCR sensitivity (cfr. Assay validation) and homogeneous risk in the target population (cfr. Geographical sampling).

The structure of the model is as follows (4):

$$\begin{aligned} p &\sim Beta(\alpha, \beta) \\ sens_{swab} &\sim Beta(\alpha_{swab}, \beta_{swab}) \\ p_{test} &\sim 1 - Binomial(0, s, p \times sens_{swab}) \\ k &\sim Binomial(w, p_{test}) \end{aligned}$$

The method models the prevalence p in the sample as Beta distributed with shape parameters α, β, and uses draws from the distribution to simulate the test results for a pool with s individual samples in it.

The rate $p_{test}$of positive pools k out of w total tested pools is modelled as a binomial distribution. The swabbing $sens_{swab}$ is modeled as a Beta distribution with shape parameters $\alpha_{swab}, \beta_{swab}$.

Bayesian prior definition

The Bayesian models permit the definition of priors on the prevalence and the swabbing sensitivity, incorporating a-priori expectations on the possible range of results, making them more robust against noise due to small sample sizes. We suggest a default value of 0.3 for α, β, which indicate large uncertainty regarding the real prevalence, but also favours lower values, in accordance with results from a number of studies showing prevalence figures constantly below 5-10% (see background), and also to correct the tendency of pooled analysis to overestimate prevalence [30] when underpowered. For $\alpha_{swab}, \beta_{swab}$, we suggest respectively 8.8 and 0.74, which represent a (median) swabbing sensitivity of ~95% ranging from 70% to 100%.

Setting custom priors for the prevalence can be useful in two settings:

- Isolated point-prevalence studies with a limited number of samples available and relatively strong certainty around the range of possible values for the prevalence, e.g. in institutes where random or census studies have been done in similar settings;
- Periodic studies, repeated in the same location, in which the results of the previous iteration can be used to inform the next one. This methodology is useful to perform multiple studies over time while keeping the number of required samples low. Priors should not be too tight in order to not obscure trends in the prevalence;

To evaluate priors for the prevalence, a value u must be chosen which indicates the intended precision, with higher values indicating more certainty (tighter 95%CrI). To have a distribution whose mean prevalence is p_{prior}, α must be set to $u \times p_{prior}$ and $\beta = u(1 - p_{prior})$; if instead p_{prior} is intended to be the most probable value (mode), then $\alpha = p_{prior}(u - 2) + 1$ and $\beta = (1 - p_{prior})(u - 2) + 1$. In the case of the conjugate Bayesian method, the priors need to be put on the pool test result probability (p_{test}), not directly on the prevalence (p). In the PooledPrevalence R Package associated to this document, we provide tools to simplify the selection of α, β in order to set the appropriate prior distribution. It is important to specify wide enough priors not to suppress the information coming from the data but which also discourage unrealistic results due to chance, especially when the sample size is low. Results should be reported with and without the use of custom priors, to assess the impact on the estimates.

Optimisation of test design

It is important to implement a study design that enables prevalence estimates with good accuracy with a limited use of resources. The goal is to achieve a low estimation error (bias) and uncertainty (variance) while keeping the total number of tested individuals and consumption of laboratory material low.

The optimisation of a pooled testing study is a complex process, since it involves finding a balance between the pool size (number of individual samples in a testing pool), which increases bias but decreases costs, and the number of pools and the total number of samples, which decreases variance at the cost of more resources. We developed a tool in the PooledTesting R package to help identify optimal study design. The tool requires user defined operational limits (maximum pool size, maximum number of pools to test, maximum number of individual samples) and a hypothetical prevalence of the disease in the target population. The tool uses the following algorithm:

1) The optimisation algorithm computes the expected number of positive pools in a study, given a hypothesised prevalence p_{hyp}, for each combination of pool size and number of pools. The range of evaluated combinations depends on the operational limits chosen by the user in terms of maximum pool size, number of pools to test and number of individuals samples.

2) From the simulated results, the prevalence is back-estimated using the conjugate Bayesian method defined in (3) and the estimation error $err = abs(p - p_{hyp})$ and the uncertainty $unc = p_{97.5\%} - p_{2.5\%}$ are computed together with the equivalent estimated err_{base}, unc_{base} obtained by an unpooled analysis (with one sample per pool and number of tested pools equal to the number of individuals in the sample).

3) Finally a cost function is estimated for each combination: $cost\ score = \log\left(\frac{err}{p_{hyp}}\right) + \log\left(\frac{unc}{p_{hyp}}\right) + \log\left(\frac{err_{base}}{p_{hyp}}\right) + \log\left(\frac{unc_{base}}{p_{hyp}}\right) + \log(\frac{w}{s})$. The cost score decreases as the statistical accuracy (related to the first 4 terms in the equation) increases but grows with the resource usage represented by $\frac{w}{s}$. To keep the cost down the pool size s needs to be increased or the total number of must be limited.

4) A decision tree algorithm identifies relevant combinations of number of tested pools and pool size, defining a number of 'optimisation windows' that minimise the cost function. The windows are proposed to the user starting by those with the best balance between resource utilisation and estimation accuracy.

Here we provide an example of an output of the optimisation algorithm for a hypothesised prevalence of 0.5%. The tool was also given as an input the following operational limits: maximum pool size = 12, maximum number of pools to test = 2 000, and maximum individual samples = 2 000. The maximum number of individual samples act as an upper limit to the possible combinations of pool size and number of pools: the product of the two needs to be lower than the number of individual samples; for example, a study design with 500 tested pools of size 12 (6 000 individuals) is not considered given the maximum of 2 000 individual samples.

With these inputs, the best suggested 'optimisation window' has a pool size ≥ 9 and a number of tested pools ≥ 54 (table 1, first column). To account for possible problems of sensitivity as the pool size increases, we suggest choosing the lower limit (pool size = 9) of the window. Regarding the number of pools, these results indicate that using just 54 pools is enough to achieve good accuracy; nevertheless, we suggest testing the largest number of pools possible for the given the operational limits, in this case 2000/9 = ~222 (table 1, second column). In table 1 we report the simulated results for both possibilities. The results represent the median and the [2.5%, 97.5%] interval of the distribution of the estimates of 2000 simulations.

Table 1. Simulations of study results given a prevalence of 0.5%, and the best operational windows returned by the optimisation algorithm for the operational limits described in the text

	w = 54, s = 9	w = 222, s = 9
Saved tests (%)	432 tests (88.9%)	1 778 tests (88.9%)
Total tested individuals	486	2 000
Estimated prevalence	0.45%, 95%CI: [0.047%, 1.3%]	0.52%, 95%CI: [0.21%, 0.84%]
Estimation uncertainty	1.2%, 95%CI: [0.51%, 2.1%]	0.64%, 95%CI: [0.41%, 0.81%]
Estimation error	0.26%, 95%CI: [0.047%, 0.83%]	0.12%, 95%CI: [0.019%, 0.34%]
Unpooled study prevalence	0.45%, 95%CI: [0.047%, 1.3%]	0.51%, 95%CI: [0.21%, 0.86%]
Unpooled study uncertainty	1.2%, 95%CI: [0.52%, 2%]	0.63%, 95%CI: [0.41%, 0.81%]
Unpooled study error	0.26%, 95%CI: [0.053%, 0.77%]	0.11%, 95%CI: [<0.001%, 0.34%]

The loss in accuracy between pooled (s = 9) and unpooled (s = 1) study designs, given the same number of tested individuals is negligible but the pooling leads to a significant saving of resources (88.9% less tests used). Further, performing 222 tests instead of 54 almost doubles the accuracy. Given the lower number of tests required after pooling, the extra resources should be used to increase the total number of tested individuals, since that greatly decreases estimation uncertainty and error.

Table 2 shows the results of when the algorithm is applied in a scenario with a higher hypothesised prevalence (5%), keeping the same operational limits (maximum pool size = 12, maximum pools to test = 2 000, maximum tested individuals = 2 000). In this case, the algorithm suggests a window of $s \geq 8$ and $w \geq 102$. Again we show the results for the most conservative scenario (w = 102, s = 8) and after maximisation of the number of tests for the same pool size (w = 250, s = 8). The results represent the median and the [2.5%, 97.5%] interval of the distribution of the estimates of 2000 simulations.

Table 2. Simulations of study results given a prevalence of 5%, and the best operational windows returned by the optimisation algorithm for the operational limits described in the text

	w = 102, s = 8	w = 250, s = 8
Saved tests (%)	714 tests (87.5%)	1 750 tests (87.5%)
Total tested individuals	816	2 000
Estimated prevalence	5%, 95%CI: [3.6%, 6.8%]	5%, 95%CI: [4%, 6.1%]
Estimation uncertainty	3.3%, 95%CI: [2.7%, 3.9%]	2.1%, 95%CI: [1.9%, 2.4%]
Estimation error	0.56%, 95%CI: [0.048%, 2%]	0.36%, 95%CI: [<0.001%, 1.2%]
Unpooled study prevalence	5%, 95%CI: [3.6%, 6.5%]	5%, 95%CI: [4.1%, 6%]
Unpooled study uncertainty	3%, 95%CI: [2.5%, 3.4%]	1.9%, 95%CI: [1.7%, 2.1%]
Unpooled study error	0.53%, 95%CI: [0.043%, 1.8%]	0.36%, 95%CI: [<0.001%, 1.1%]

The pooled estimation loses efficiency as the prevalence grows; so to keep good accuracy, a larger number of tested pools and a lower number of individuals per pool is required. As a rule of thumb, it is better to choose a hypothesised prevalence slightly higher than what is expected, to avoid the risk of an underpowered study in case the initial estimates were too conservative.

These results show how it is possible to save a large amount of resources using a pooled testing strategy. Even if the uncertainty in the unpooled test is expectedly lower, the difference is trivial, especially compared to the number of tests saved. The tool we provide allows study organisers to choose among a number of settings based on their specific situation in terms of resource constraints and disease prevalence.

Conclusion

We presented a methodology to obtain prevalence estimates for SARS-CoV-2 infections based on PCR testing when pooling is performed to improve the efficiency of the use of laboratory resources. Member States may consider starting to perform such studies, e.g. based on samples of 1 000 individuals by subnational region. When repeating this at e.g. 2-3 week intervals these estimates may provide a valuable complement to the other surveillance systems. Through a Bayesian approach, the results from previous studies may feed into improving the estimates of following studies. We provide an R package (https://github.com/EU-ECDC/PooledPrevalence) to assist Member States in obtaining their prevalence estimates.

Consulted experts (in alphabetical order)

ECDC experts: Cornelia Adlhoch, Eeva Broberg, Bruno Ciancio, Angelo D'Ambrosio, Joana Gomes Dias, Aftab Jasir, Helen Johnson, Irina Jovel Quinonez Dalmau, Pete Kinross, Katrin Leitmeyer, Angeliki Melidou, Pasi Penttinen, Amelie Plymoth, Chantal Quinten, Emmanuel Robesyn, Gianfranco Spiteri

External experts: Cyril Barbezange (Sciensano, Belgium), Jessica Beser (Public Health Agency of Sweden, Sweden), Ben Cooper (University of Oxford, United Kingdom), Ramona Groenheit (Public Health Agency of Sweden, Sweden), Rene Niehus (Harvard T.H. Chan School of Public Health, USA), Kristjan Vassil (University of Tartu, Estonia)

References

1. European Centre for Disease Prevention and Control (ECDC). Coronavirus disease 2019 (COVID-19) in the EU/EEA and the UK - ninth update [updated 23/04/2020]. Available from: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-rapid-risk-assessment-coronavirus-disease-2019-ninth-update-23-april-2020.pdf.
2. Lavezzo E, Franchin E, Dannenburg GC, Ciavarella C, Barzon L, Del Vecchio C et al. Suppression of COVID-19 outbreak in the municipality of Vo, Italy. MedRxiv. 18/04/2020.
3. Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic Population. N Engl J Med. 2020 Apr 14.
4. Folkhälsomyndigheten. Förekomsten av covid-19 i region Stockholm, 26 mars–3 april 2020. 20/04/20. Available from: https://www.folkhalsomyndigheten.se/contentassets/7bd5627f82a84590bc2992784234b88b/forekomsten-covid-19-region-stockholm-26-mars3-april-2020.pdf.
5. Folkhälsomyndigheten. Swedish national PCR prevalence study. 12/05/20. Available from: https://www.folkhalsomyndigheten.se/nyheter-och-press/nyhetsarkiv/2020/maj/nya-resultat-fran-undersokning-av-forekomsten-av-covid-19-i-sverige/.
6. Office of National Statistics. Coronavirus (COVID-19) Infection Survey pilot: England. 14/05/2020. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/england21may2020.
7. Lakshmi N. Theagarajan. Group Testing for COVID-19: How to Stop Worrying and Test More. arXiv. 14/04/2020.
8. Itsik Bergel. Variable pool testing for infection spread estimation. arXiv. 07/04/2020.
9. Dorfman R. The detection of defective members of large populations. The Annals of Mathematical Statistics 14.4 (1943): 436-440.
10. M. Sobel PAG. Group testing to eliminate efficiently all defectives in a binomial sample. Bell Labs Technical Journal.38(5):1179–1252, 1959.
11. Morris U, Xu W, Msellem MI, Schwartz A, Abass A, Shakely D, et al. Characterising temporal trends in asymptomatic Plasmodium infections and transporter polymorphisms during transition from high to low transmission in Zanzibar, 2005-2013. Infect Genet Evol. 2015 Jul;33:110-7.
12. Ngondi JM, Ishengoma DS, Doctor SM, Thwai KL, Keeler C, Mkude S, et al. Surveillance for sulfadoxine-pyrimethamine resistant malaria parasites in the Lake and Southern Zones, Tanzania, using pooling and next-generation sequencing. Malar J. 2017 Jun 5;16(1):236.
13. Nasa Sinnott-Armstrong DK, Klein D, Hickey B,. Evaluation of Group Testing for SARS-CoV-2 RNA. 30/03/2020.
14. Christian Gollier and Olivier Gossner. Group testing against Covid-19. Covid Economics vol. 1, n. 2, April 2020, pp. 32–42.
15. Hogan CA, Sahoo MK, Pinsky BA. Sample Pooling as a Strategy to Detect Community Transmission of SARS-CoV-2. JAMA. 2020;323(19):1967-9.
16. Shental N, Levy S, Skorniakov S, Wuvshet V, Shemer-Avni Y, Porgador A, et al. Efficient high throughput SARS-CoV-2 testing to detect asymptomatic carriers. medRxiv. 20/04/2020.
17. European Centre for Disease Prevention and Control (ECDC). Surveillance of COVID-19 at longterm care facilities in the EU/EEA. 2020 19/05/2020.
18. Plymoth A, Rotzen-Ostlund M, Zweygberg-Wirgart B, Sundin CG, Ploner A, Nyren O, et al. Self-sampling for analysis of respiratory viruses in a large-scale epidemiological study in Sweden. Euro Surveill. 2015 Mar 19;20(11).
19. European Centre for Disease Prevention and Control (ECDC). Laboratory support for COVID-19 in the EU/EEA 2020 [22/05/2020]. Available from: https://www.ecdc.europa.eu/en/novel-coronavirus/laboratory-support.
20. Azzi L, Carcano G, Gianfagna F, Grossi P, Gasperina DD, Genoni A, et al. Saliva is a reliable tool to detect SARS-CoV-2. Journal of Infection. 2020 2020/04/14/.
21. Mawaddah Azman HSG, Sai Guan Lum, Marina Mat Baki,. Upper respiratory tract sampling in COVID-19. Malaysian J Pathol 2020; 42(1) : 23 – 35.
22. Wyllie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M, Vijayakumar P, et al. Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs. medRxiv. 2020:2020.04.16.20067835.
23. Abdalhamid B, Bilder CR, McCutchen EL, Hinrichs SH, Koepsell SA, Iwen PC. Assessment of Specimen Pooling to Conserve SARS CoV-2 Testing Resources. Am J Clin Pathol. 2020 May 5;153(6):715-8.
24. Lohse S, Pfuhl T, Berkó-Göttel B, Rissland J, Geißler T, Gärtner B, et al. Pooling of samples for testing for SARS-CoV-2 in asymptomatic people. The Lancet Infectious Diseases.
25. Yelin I, Aharony N, Shaer-Tamar E, Argoetti A, Messer E, Berenbaum D, et al. Evaluation of COVID-19 RT-qPCR test in multi-sample pools. medRxiv. 2020.
26. Swallow William H. Group Testing for Estimating Infection Rates and Probabilities of Disease Transmission. Phytopathology 1985;75:882-889.
27. Chick Stephen E. Bayesian models for limiting dilution assay and group test data. Biometrics 1996;1055-1062.
28. Huang Shih-Hao ea. Optimal group testing designs for estimating prevalence with uncertain testing errors. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 79.5 (2017): 1547-1563.
29. Bilder Christopher R. et. al. binGroup: a package for group testing. The R journal 2.2 (2010): 56.
30. Nicholas A. Pritchard and Joshua M. Tebbs. Estimating Disease Prevalence Using Inverse Binomial Pooled Testing. Journal of Agricultural, Biological, and Environmental Statistics 23/06/2010;Volume 16, pages70–87(2011).