Pool Strategy for Surveillance Testing of SARS-CoV-2

Felipe Marceca[1,2], Luciana Rocha-Viegas[1,3], Nicolás Pregi[1,4], Gabriela Barbas[5], Daniela Hozbor[6], Adali Pecci[1,3] and Roberto Etchenique[1,7,*]

[1]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Buenos Aires, Argentina
[2]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Matemática, Instituto de Investigaciones Matemáticas "Luis Santaló", CONICET, Buenos Aires, Argentina
[3]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Instituto de Fisiología, Biología Molecular y Neurociencias, CONICET, Buenos Aires, Argentina
[4]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Instituto de Química Biológica de la Facultad de Ciencias Exactas y Naturales, CONICET, Buenos Aires, Argentina
[5]Secretaria de Prevención y Promoción de la Salud, Ministerio de Salud de la Provincia de Córdoba
[6]Instituto de Biotecnología y Biología Molecular, Facultad de Ciencias Exactas, Universidad Nacional de La Plata y CCT La Plata, CONICET, La Plata, Argentina
[7]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Química Inorgánica, Analítica y Química Física, Instituto de Química Física de los Materiales, Medio Ambiente y Energía, CONICET, Buenos Aires, Argentina

* Corresponding author. E-mail: rober@qi.fcen.uba.ar

Abstract

Due to the great morbidity and mortality in the risk groups of the pandemic COVID-19 caused by the emerging coronavirus SARS-CoV-2 and in the absence of effective therapeutic or preventive measures, quarantines, social distancing and the use of masks were the measures most used by health systems to reduce infections. The social, economic and health impact caused by these measures have begun to be evaluated in the different countries. These analyses lead to underestimations because in general they evaluate disease confirmed by a laboratory test and in some cases by epidemiological link without considering asymptomatic or oligosymptomatic infection. Therefore, mitigating fast circulation of the virus requires continuous tracking, detection, and isolation of cases, for which active surveillance able to address asymptomatic cases can make a valuable contribution over the dynamics of the disease in a given society, and to allocate adequate health resources and evaluate the effectiveness of control measures. Mathematical models such as the Susceptible-Exposed-Infectious-Removed (SEIR) allow not only to improve the estimates of the evolution of the pandemic at the local level, but also to evaluate health strategies. In the context of large testing requirements and the expansion of such testing capacity, it is also essential to develop approaches that improve the efficient use of these resources. Active surveillance undoubtedly contributes to improving estimates of virus circulation and it is of particular importance in vulnerable groups of high population density that have one or more risk factors, difficult access to the health system, and inhabit semi-closed facilities such as residential care homes, mental hospitals, prison houses, police stations housing prisoners, etc. Group testing strategies are especially useful for routine community survey and for monitoring of cohesive groups. While the frequency of infection in a population, who have only some symptoms compatible with the disease or do not have any symptoms, may be low, diagnosing even a single positive person typically requires quarantine of the entire group to prevent further spread in the community. In these surveillance strategies, pooling may allow more routine monitoring and detection of low frequency of carriage, thereby improving estimates, informing policy makers, reducing transmission, and alleviating the strain on healthcare services. By means of molecular tests based on RT-qPCR, the pooling strategy has been assayed with different algorithms also for COVID-19, particularly in the asymptomatic population, since a low prevalence of the disease is expected there. This has increased SARS-CoV-2 testing throughput while maintaining high sensitivity. Here, we discuss the relevance of some active surveillance strategies to determine key facts about COVID-19 pandemics and review different testing strategies that different countries have applied for tracking SARS-CoV-2.

Keywords: SARS-CoV-2, COVID-19, coronavirus, pooling, active surveillance

Introduction

At the end of last year, China informed the United Nations (UN) health agency of the existence of several cases of atypical pneumonia in the city of Wuhan caused by a novel coronavirus (SARS-CoV-2). On January 30, when there were only 80 cases outside of China, the UN declared COVID-19 an international public health emergency, the highest alert that the UN health agency can issue to warn governments about the need to take measures to prevent its spread.

At the beginning of February 2020, the World Health Organization (WHO) convenes a research and innovation forum on COVID-19, attended by more than 400 experts and funders from around the world. By March 11, due to the alarming levels of spread and severity of the disease, the WHO determined in its assessment that COVID-19 is a pandemic. The recently emerged coronavirus pandemic is critically challenging healthcare systems around the world. As of March 17, the WHO reported more than 170,000 cases with more than 10,000 new diagnoses added in those last 24 hours; and till September 2020 SARS-CoV-2 has infected more than 25 million people and provoked more than 840,000 deaths all over the world. Nowadays (March, the 24th 2021), nearly 12902242 cases of COVID-19 have been reported, including more than 2,5 million deaths with several cities going through their second wave of increased cases (https://covid19.who.int/). In Argentina, particularly in the Province of Buenos Aires, the first confirmed case was detected on March 8, three days later than the first detected case in the country. From this date on, the rate of increase in the weekly average of total number of cases as well as in the daily reported cases was fast until the end of May, although few cases were still reported.

The complex COVID-19 symptomatology was recently classified in six groups that might correlate with illness severity [1,2]. Elderly, and those with underlying medical conditions like heart and lung problems, high blood pressure, diabetes, or cancer, are at higher risk of developing serious illness. Meanwhile, other age groups can become infected and develop moderate symptoms or even carry the SARS-CoV-2 infection asymptomatically. These groups of oligosymptomatic, presymptomatic and asymptomatic people represent a great concern for the health system since they may go unnoticed while contributing to the circulation of SARS-CoV-2 [3–5].

The pandemic context more clearly shows that urgent need for systematic collection, analysis and interpretation of health data are fundamental for the planning, implementation and evaluation of public health policies. It is important to remember at this point that epidemiologic surveillance in general, and for COVID-19 in particular, is supported by passive surveillance systems, active surveillance and specialized or sentinel systems. The first begins when the case consults the health personnel and is registered through the usual case notification system. In passive surveillance, data is collected from as many sources as possible. In active surveillance, the personnel in charge actively seek information about the disease under investigation. Medical personnel are contacted, health care centers are visited, health records are analyzed for signs of the disease, and cases are searched in the field. If suspicious cases are detected, samples are taken and sent to laboratories for analysis. All actors who can offer information on possible cases, with which they have been in contact with, are contacted. As part of active surveillance, the competent authorities are also quickly informed through previously established channels. The sentinel surveillance system uses high-quality data, collected in specialized centers that are carefully selected. Centers and professionals specialized in the disease under surveillance and high-quality diagnostic laboratories generally participate.

The first population study collecting information on the demography, clinical presentation, hospitalization, contact network and the presence of SARS-CoV-2 infection, was carried out during February 2020 in Padua (Italy). In this study, two surveys were conducted, with a time difference of two weeks. A prevalence of infection of 2.6% was found in the first survey, which was carried out at the beginning of the town lockdown. In the second survey, which was carried out at the end of the confinement, prevalence down to 1.2% [6], supporting the efficacy of the control measures. A second study, performed in Iceland, during March 2020 compared two sampling strategies, one using an open invitation method to the entire population of Iceland, and a second sampling on a random population. [7]. Then, another study performed by the Public Health Agency of Sweden, in collaboration with Swedish Armed Forces, conducted a prevalence survey in Stockholm. In this study, 2.5% tested positive for SARS-CoV-2 infection, estimating 59364 infected people in Stockholm in the last week of March. This protocol was repeated at end of April and resulted in a prevalence of 2.3%. However, a nation-wide study revealed a prevalence of 0.9% (https://www.folkhalsomyndigheten.se/the-public-health-agency-of-sweden/). Finally, 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 at the UK from April 27 to May 10 that estimated an average of 0.27% to be infected in the community population (https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/england21may2020).

All these studies were based on individual testing, which makes them expensive and time consuming to reproduce on large scales. In this way, conducting large-scale population RT-qPCR studies through pooling samples seems to be the most efficient strategy to perform.

Regardless of the surveillance system used, the novelty of the disease made it necessary from the beginning to have a specific laboratory tool for detecting SARS-CoV-2 infection in order to confirm the clinical case. This methodology refers to molecular techniques based fundamentally on reverse transcription techniques and quantitative polymerase chain reaction (PCR). Depending on the objective, this strategy can be applied to a sample from a single individual or to a pool of samples from different individuals. This last procedure will be our main focus, since it is used to gain information at a population level and is essential for the continuous tracking, detection and isolation of cases.

The pooling of individual samples as a cost-saving method for diagnosis of infectious diseases is presented as a possible strategy to confront the COVID-19 pandemic. In the past, this group testing strategy was successfully used for identifying individuals with syphilis, HIV, hepatitis B virus, hepatitis C virus, *Chlamydia trachomatis* and *Neisseria gonorrhoeae* using different biological samples [8–15]. As for COVID-19, the pooling strategy has been applied for the general population in different countries *i.e* Australia [16], China [17], Ecuador [18], Germany [19], Ghana [20], India [21], Israel [22], Malaysia [23], Malawi [20], South Africa [20], Tunisia [24], Uruguay [25], US [26,27] and Zimbabwe [20] among others.

In accordance to the US Centers for Disease Control and Prevention (CDC) guidance, laboratories certified under the Clinical Laboratory Improvement Amendments (CLIA) can use a specimen pooling strategy to expand SARS-CoV-2 nucleic acid diagnostics or screening testing capacity when using a test authorized for such use by the U.S. Food and Drug Administration (FDA) [Policy for Coronavirus Disease-2019 Tests During the Public Health Emergency (Revised) ("Policy for COVID-19 Tests")]. In other countries, this strategy is carried out in experienced laboratories that work in coordination with the ministries or secretaries of health. In its simplest form, this strategy consists of two-stages: first, mixing several samples together in a single tube and then testing the pooled samples in only one reaction. If the pooled sample is negative, it can be presumed that all individuals were negative. However, if the pooled sample turns out positive and individual identification is requested, a second stage is needed, where each sample will be separately tested to find out which individual sample was responsible for that result. In low prevalence scenarios, because pooled samples are often negative, ultimately fewer tests are run overall, saving time and testing supplies, and therefore, the results can be returned to patients more quickly in most cases. The advantages of this two-stage specimen pooling strategy include preserving testing reagents and resources, reducing the amount of time required to test large numbers of specimens, and lowering the overall cost of testing.

By implementing the eight-sample Dorfman pooling to test 26,576 samples from asymptomatic employees from the Hadassah Medical Centre (HMC) of Israel, 31 (0.12%) SARS-CoV-2 positives were identified, thus achieving a 7.3-fold increase in throughput [28]. In another report, by using a Shiny application (https://www.chrisbilder.com/shiny), it was proposed that the most efficient pool size is five specimens [27]. In Argentina, this strategy was performed successfully in semi-closed facilities in the Province of Buenos Aires at different pandemic periods [29]. With this strategy, the efficiency, measured as people tested per determination, ranged from 2 to 7.4, depending on the place of origin, with an average of 3. Remarkably, it allowed the early detection of outbreaks, and in addition, has evidenced that infected individuals, either asymptomatic or presymptomatic, may present cycle threshold (Ct) values as low as those of symptomatics.

Here we review those group testing strategies that have been applied for tracking COVID-19, the different algorithms used for saving resources compared to testing samples individually and the assessment of pooling performance. We also discuss the importance of group testing to carry on an active surveillance able to address a valuable contribution for the detection of asymptomatic cases.

Molecular diagnostic tests based on RT-qPCR

Since the beginning of the pandemic, there was a desperate need to develop diagnostic tests for COVID-19 to limit the spread of the virus. The whole genome sequencing of SARS-CoV-2 performed on January 10 led scientists to design testing protocols to detect the pathogen in the affected people. Though alternative diagnostic techniques to use at the point of care are also required, real time reverse transcription polymerase chain reaction (RT-qPCR)-based methodologies have been widely recommended. In fact, the WHO proposed this technique as the frontline diagnostic approach to detect SARS-CoV-2 infection in suspected patients. The assay is carried out by isolating RNA from a clinical sample (nasopharyngeal or oropharyngeal swabs, saliva) and adding it to a qPCR master mix containing specific forward and reverse primers.

Most countries, the US and Argentina among them, mainly use a one-step PCR format to diagnose COVID-19. The list of the different kits approved for use in these countries are detailed on these web pages https://www.fda.gov/media/134922/download, https://www.argentina.gob.ar/anmat/regulados/productos-medicos/reactivos-covid-19, respectively.

Group testing strategies

As mentioned before, the COVID-19 pandemic has attracted renewed attention to group testing strategies. By using a single test for a group of samples that are pooled together, these strategies allow the identification of positive samples efficiently, saving time and resources.

Group testing is any procedure that allows identifying certain specimens by testing groups rather than each specimen individually. It was first introduced by Dorfman to dramatically reduce the number of tests needed to screen future soldiers for syphilis [30]. Since then, it has grown into a rich subject that can be studied from different perspectives such as statistics, combinatorial mathematics and information theory. Applications include the identification of positive samples in medicine which will be our main focus, physical mapping of genes in molecular biology, multiple access communications, data storage and compression, and cybersecurity [31–33]. Let us describe the main idea behind group testing. For a set of samples that can either be positive or negative one can group any number of them together in a pool and perform a single test. If the test is negative every sample in the pool is presumed negative, whereas if the test is positive there is at least one positive sample in the pool. For example, Dorfman's original algorithm consists of two stages. First the samples are divided in groups of a certain size and each group of samples is tested in a single pool. During the second stage, samples in positive pools are tested individually and all positive samples are identified (Figure 1).
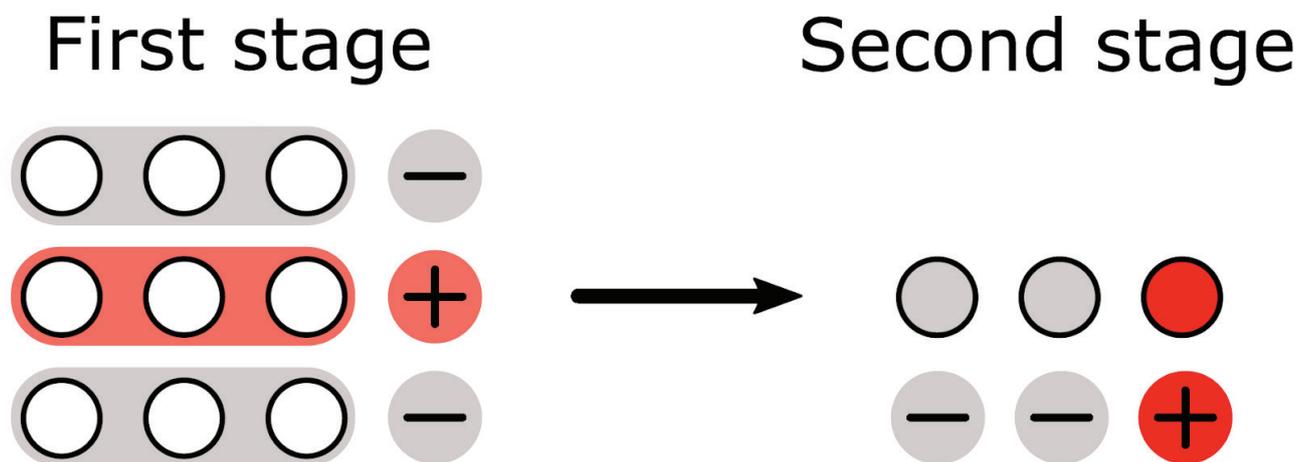

Figure 1. Scheme of Dorfman's algorithm: Two stages are carried out. Samples are divided into groups and for each group, one test is made pooling its samples together. Then, samples from positive pools are tested individually.

When the number of positive samples is small, group testing allows to quickly rule out large numbers of negative samples using few tests. Group testing strategies delve into the optimal design of pooling schemes according to different criteria such as minimizing number of tests, time or other resources. They are also subject to constraints such as maximal pool size due to dilution effects [22]overwhelmed healthcare systems are already experiencing shortages of reagents associated with this test, calling for a lean immediately applicable protocol. Methods: RNA extracts of positive samples were tested for the presence of SARS-CoV-2 using reverse transcription quantitative polymerase chain reaction, alone or in pools of different sizes (2-, 4-, 8-, 16-, 32-, and 64-sample pools, maximum number of simultaneous tests due to equipment capacity, etc. For a given group-testing method, parameters such as the sizes of the pools are optimized depending on the presumed positivity (the proportion of infected individuals in the whole sample). If the individuals are randomly chosen from the population, it equals the prevalence. The lower the prevalence, the bigger the pool-sizes and the bigger the savings on tests will be. For example, to achieve the minimum number of expected tests in Dorfman's method for a prevalence p one has to use a pool of size roughly $1/\sqrt{p}+1/2$ and one should expect to reduce the number of tests by a factor of $2\sqrt{p}$. Therefore, it is important to know how sensitive the method is to changes in prevalence and to have a good estimation of this parameter (Figure 2). Also notice that pooling strategies are more effective than individual testing only when the prevalence is sufficiently small. For example, Dorfman's method offers no improvement whenever $p\approx0.25$ (that is, p≈25%) so that $2\sqrt{p}\approx1$. It should be mentioned that usually these models assume samples are independent. This is not the case when testing closed facilities such as nursing homes, where samples have a high correlation that should be taken into account in the design of the pooling scheme.
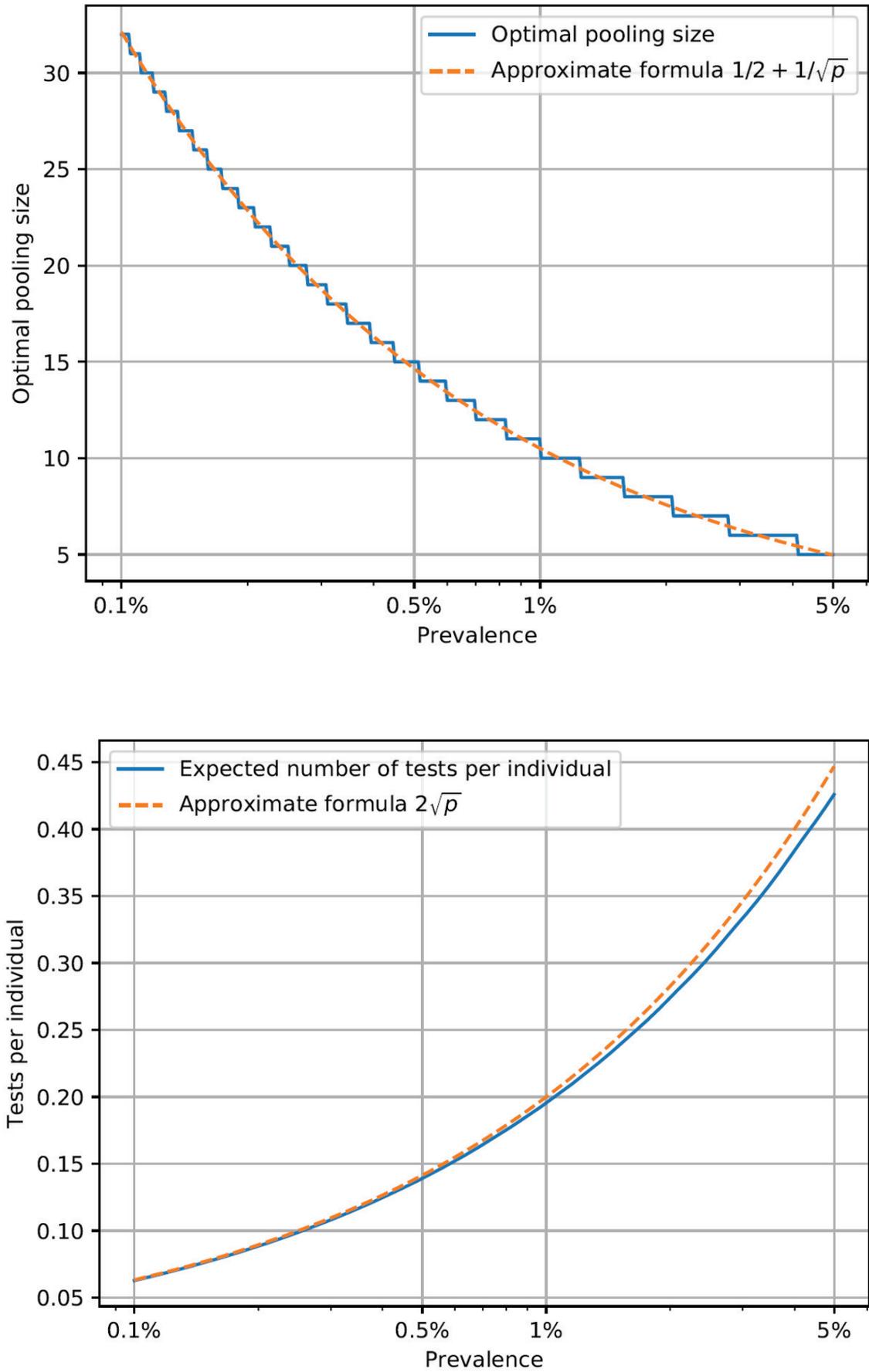
Figure 2. *Upper panel:* optimal pooling size for Dorfman's algorithm in terms of the prevalence that minimizes the expected number of tests per individual. *Lower panel:* expected number of tests per individual for Dorfman's algorithm in terms of the prevalence when using the optimal pooling size.

There are two general types of group testing schemes known as adaptive and non-adaptive. Adaptive strategies perform tests sequentially and use the results of previous tests to determine which test will follow. This makes them more effective in saving tests and easier to tailor but may consume more time and be more difficult to scale up. For non-adaptive strategies on the other hand, tests are predetermined and conducted simultaneously. This allows streamlining test-flow and avoiding waiting for previous results to prepare the next batch of tests. However, non-adaptive approaches may sometimes leave samples unclassified if the proportion of positive samples is unusually high.

Dorfman's original algorithm is a simple example of an adaptive scheme that is being implemented in several countries to identify SARS-CoV-2 among asymptomatic individuals (see for example [28]). There are also several variants of Dorfman's method which have been discussed in the context of the COVID-19 pandemic. One approach known as nested pooling, is to increase the number of stages by dividing positive pools into smaller sub-pools and iterate this process a specified number of times before reaching individual testing [34,35] (Figure 3). Increasing the number of stages can notably reduce the number of tests required (even among other group-testing strategies), but also adds time and complexity to the procedure.
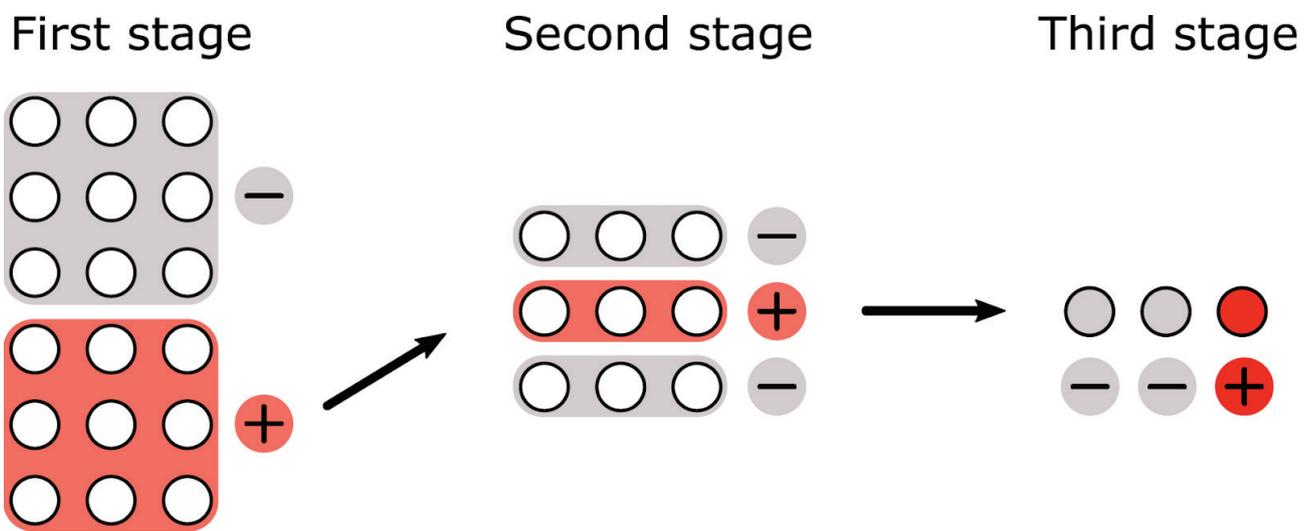

Figure 3. Nested pooling. Iteration process based on dividing positive pools into smaller sub-pools a specified number of times before reaching individual testing.

Another approach is to maintain only one group-testing stage and one individual-testing stage as in Dorfman's method, but to improve it by implementing a more sophisticated group-testing design. This can be done for example by arranging samples in a square matrix and testing each row and each column in a different pool (Figure 4). In other words, each sample is included in two pools, one for its row and one for its column. Since samples on negative pools are considered to be negative, only samples whose row and column are positive have to be tested in the individual stage. In particular, if there is only one positive sample, it can be identified directly in the first stage and no further testing is needed. A suboptimal pool-size can be tuned so that this happens quite frequently prioritizing saving time over tests. This method was implemented for surveillance in closed facilities such as nursing homes [29]. Generalizations replacing matrices with higher dimensional arrays were proposed [36].
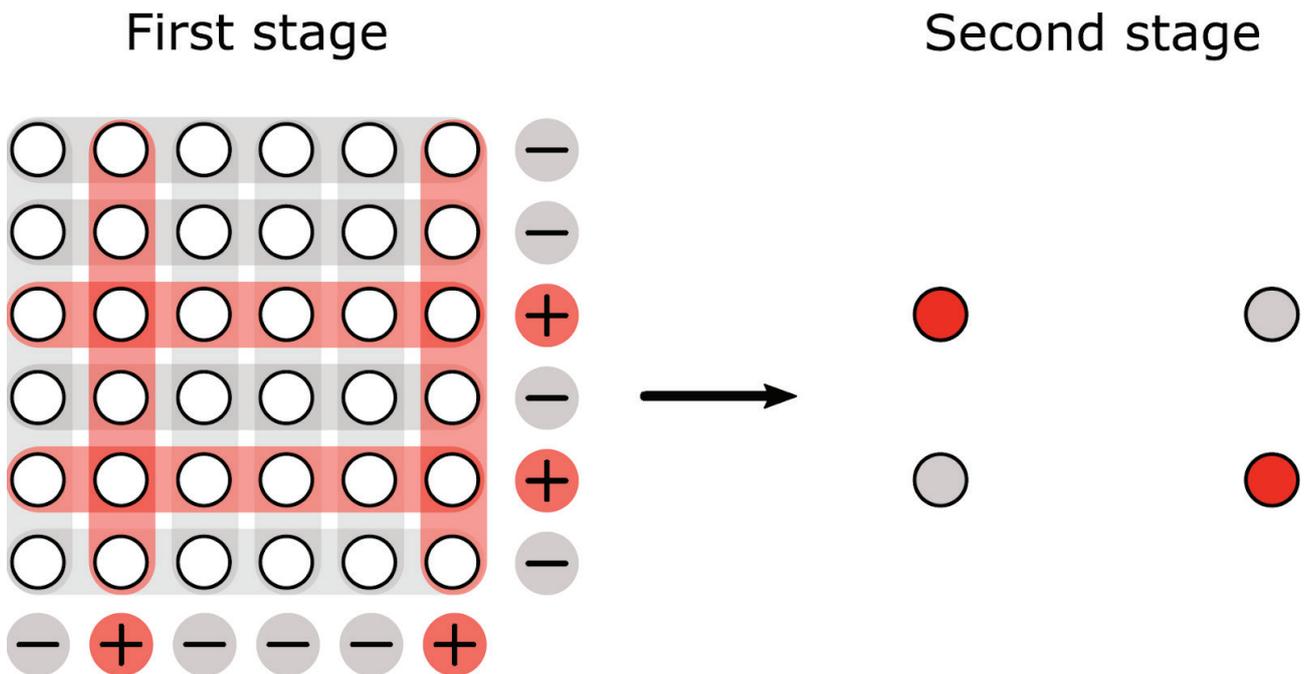
First stage

Second stage

Figure 4. Matrix pooling. Square-matrix arrangement where each sample is included in two pools, one for its row and one for its column. Only samples whose row and column are positive are then tested in the individual stage.

Regarding non-adaptive methods, a single group-testing stage is designed including each sample in several pools. The assignment is done in such a way that the maximum number of positive samples can be identified with the least number of pools in one go. These designs are usually inspired on objects well-known in the context of combinatorial mathematics as well as on error correcting codes from computer sciences. For example, Steiner systems [37] and an error correcting code [38] were employed. Furthermore, in the latter, their non-adaptive pooling design was implemented to screen 1115 health care workers using 144 tests. The success of non-adaptive methods relies heavily on the proportion of positive samples not exceeding a threshold value. For this reason, prevalence has to be carefully estimated leaving a safe margin. Another drawback is that pool designs of non-adaptive methods are very intricate by nature, which may make them difficult to implement without automation.

Assessment of pooling strategy performance

Generally, the FDA recommends validating the test with either pooling approach (sample/media pooling or swab pooling) in a way that preserves the sensitivity of the test as much as possible. That is, it is preferable to use an approach where all specimens identified as positive when tested individually are also identified as positive when tested using the pooled testing approach. However, a decrease in performance is likely with pooling strategies, due to dilution of the primary clinical sample. Nevertheless, since sample pooling will greatly increase the number of individuals that can be tested using existing resources, a small reduction in sensitivity may be acceptable depending on the pooling efficiency and other mitigations in place. Therefore, the FDA generally recommends that, after pooling, test performance includes ≥85% percent positive agreement when compared with the same test performed on individual samples (https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/pooled-sample-testing-and-screening-testing-covid-19). An additional limitation is that samples from a negative pool are considered presumptive negatives but never tested individually. Therefore, the pool size must be chosen so that the amount of false negatives generated by this procedure remains low. In this sense, key principles for successful application of group testing involve knowledge of the limit-of-detection, sensitivity, and specificity of the assay, and the prevalence of disease in the population. The goal of the process is to determine a pool size that provides the greatest conservation of resources while maintaining the reliable performance of testing.

The University of Stanford in the US performed a retrospective study with pools of 9 or 10 individual samples that had not been tested for SARS-CoV-2 early in the pandemic. This pooled screening facilitated detection of early community transmission and enabled timely implementation of appropriate infection control measures to reduce spread [26]. In the meantime, within the community of Nebraska with a specimen positive rate around 5%, a pooling ratio of 1 to 5 was expected to retain accuracy and result in greater efficiency of test resources. Results indicated that all positive samples by

the non-pooled method were detected in pools with 4 other negative samples. Determination of the optimal pool size prior to testing is capable of improving the overall efficiency of pooled testing [27].

Within a 6.8% positive rate, Spanish clinical microbiology laboratories obtained excellent results in terms of sensitivity, specificity and positive and negative predictive values in pools of 9 and 10 samples [39]. However, web-based applications for pooling estimate that a pool size with 5 specimens may conserve more resources than a pool size of 10 as the prevalence rate increases [23].

In Germany, two pooling techniques were compared, a "routine high throughput" approach where random samples are pooled together for testing, or a "door to door" approach where cohesive groups (ie, families, neighbors, etc.) are pooled together for testing. While both approaches save substantial resources, the "door to door" approach was found to carry more benefit, reducing tests by 56% to 93%, whereas the "routine high throughput" resulted in 24% to 86% fewer tests. In low to moderate infection levels, even a pool size of 5 would reduce the number of tests needed by 5-fold (78%) [40].

In countries with infection levels over 20%, a pool size of 10 would still result in a considerable reduction in the number of tests required (up to 50%). These studies showed that pooled testing is able to detect positive samples with sufficient accuracy. It is best used for population-wide screening, contact tracing, and the monitoring of essential workers and asymptomatic individuals with unidentified risk, such as in airports, versus being much less effective if used in settings with high clinical suspicion, such as patients showing symptoms. While concerns exist that low positive samples, such as those found in convalescent patients, could escape detection with increasing pool size, additional amplification cycles could be employed to allow better detection of larger pools. Nevertheless, group testing shows great potential in increasing testing capacity with existing resources with minimal loss of accuracy [40].

With a positive rate (even in asymptomatic persons) within 5%, the Indian Council of Medical Research observed that pooling of 5 or 10 specimens did not affect the sensitivity of detecting SARS-CoV-2 when the PCR cycle threshold (Ct) of the original specimen was lower than 35. However, in specimens with low viral load (Ct > 35), 13.3% were false negative. Thus, in pooled samples, graphs should be analyzed for the sigmoid curve even beyond Ct value 40, and in case of the appearance of any graph, the RT-qPCR should be repeated with deconvoluted samples [41].

When prevalence is low, 5-sample pooling versus 10-sample pooling showed Ct values differences of less than 5 as compared with individual tests [23]. Even more, a single positive sample can be detected in pools of up to 16-32 samples, with an estimated false negative rate of 10%. Despite a relatively restrictive cutoff, pooling would reach sensitivity of 96% for a pool size of 16 samples. Single positive samples can be detected when pooling either after or prior to RNA extraction [22]. Consistently with this, a comparison of Ct value differences between pool sizes from 4 up to 30 samples and individual positive samples ($Ct_{pool} - Ct_{positive}$ sample) demonstrated to be in the range of up to 5 [19].

An important concept to take into account when testing by pooling is the validation of the selected analytic method. This will allow us total control over the testing. FDA suggests a plan for ongoing monitoring of the positivity rate and of the performance of a test with a pooling strategy. In this sense, laboratories must perform their own validation pool studies for kits used for each RNA extraction and DNA amplification, based on the prevalence rate of COVID-19 in their own region. Finally, pooling is an effective approach to expand the impact of limited test resources and reagents during specific stages of an infectious disease outbreak.

In the province of Buenos Aires (Argentina) pooling strategies were successfully carried out in semi-closed facilities. By comparative analysis with single tests, we detected that clusters of 5 samples for RNA extraction and 10 (two sub-pools containing RNA extracted from 5 samples) is adequate since only 2 Cts are increased on average [29]. Concordance between the 5 or 10 sample pool and individual sample testing was 100% in the Ct value ≤36 cycles. As reported by other authors, the effectiveness of sample pooling is dependent on the prevalence of positive samples within the pool [42]. The sample pooling strategy worked best in settings with low prevalence and individual subgroups with low clinical suspicion such as asymptomatic individuals. Individual groups for whom there is a higher pre-test probability and those with serious manifestations should not be included in pooled testing but should be tested individually. To date and since the end of May we have analyzed more than 11,000 samples using the pooling test strategy and the obtained data suggests that pooling of up to 5-10 samples *per* pool can increase test capacity with existing equipment and test kits and detects positive samples with sufficient diagnostic accuracy.

This strategy was also implemented by the Ministry of Health of the province of Córdoba at the Faculty of Chemical Sciences and Ceprocor Ministry of Science and Technology. From the end of July until the end of October, when the positive rate ranged from 0.7% to 6.0%, the number of people tested per test used ranged from 5 to 2.

Group testing also evidenced that asymptomatic/presymptomatic/or oligosymptomatic infected individuals may present Ct values as low as those of symptomatic individuals. Furthermore, early stages of COVID-19 outbreaks were

detected in different localities of Buenos Aires, increasing the likelihood of saving lives especially in places where risk groups were concentrated.

Importance of active surveillance to identify key facts about COVID-19 pandemics

Asymptomatic transmission of SARS-CoV-2 has been one of the most widely extended controversies since COVID-19 pandemics started. The initial evidence suggested that this new coronavirus was transmitted by respiratory droplets, mainly through coughing or sneezing as many other respiratory diseases (https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html). Additionally, surface-related contagion was considered key. As coughing is a typical symptom of COVID-19 infection, symptomatic contagion was seen as the main route by which SARS-CoV-2 enters and diffuses into societies.

However, the course of the pandemics clearly showed that other transmission ways could be responsible for most of the contagions, especially when a majority of the people assessed the risks of COVID-19 transmission and preempted any contact with other people after having compatible symptoms. Perhaps the most important of these ways is the transmission by aerosols, which acquired notoriety after the Guangzhou restaurant episode [43]. Aerosols are tiny droplets, much smaller than the ones involved in cough-related emissions. While these droplets fall down in seconds timescale, nanodrops constituents of aerosols can remain suspended for minutes [44]. This characteristic, together with the possibility of air stirring and active dissemination through convection, increases the risk of being in close spaces with an infected person, even if no symptoms are present. The mere act of talking, singing, or even breathing can transmit the virus through aerosolized particles, which are then inhaled by a susceptible person. The CDC site, after several months denying the "airborne" (aka. aerosolized) transmission, changed its view by September 2020, and stated that this mechanism constitutes a way of contagion in closed environments.

A key question to understand SARS-CoV-2 transmission is the fraction of the contagions that are due to exposure to individuals that are asymptomatic at the very moment of the contact. This definition implies that the somewhat usual differentiation between a- and pre- symptomatic must be left aside in order to focus on the practicality of the issue.

"Infodemic" about COVID-19 has profusely touched the asymptomatic issue. Its relationship with commercial activity lockdowns and travel barriers, school opening, and many other social-related situations have put this controversy in the center of political agendas. Moreover, reports of indisputable goals in the fight against COVID-19, as the massive testing after complete remission of local epidemics in Wuhan, China, have been misquoted as proof that asymptomatic contagion practically does not exist [17]. On the other side, authoritative institutions such as US CDC have put a warning on this danger, indicating that the absolute majority of the contagions are due to asymptomatic individual transmission (https://www.cdc.gov/coronavirus/2019-ncov/more/masking-science-sars-cov2.html). The transmission potential of any individual arises from two main factors: the number of close encounters with other subjects and the probability of transferring SARS-CoV-2 to the susceptible one. While the first depends mainly on social factors, the latter is closely related to the viral load in the upper respiratory tract of the infected person.

In our experience, around ten thousand of asymptomatic individuals were tested within the framework of active surveillance strategies designed by the Ministry of Health of the Province of Buenos Aires [29]. That means that we have detected hundreds of asymptomatic/presymtomatic individuals who were positive for SARS-CoV-2 RT-qPCR, and their Ct values can be used as a proxy of their viral loads and compared with symptomatic persons from the same regions.

Figure 5 shows the histograms for the N gene Ct values in two sets of clinical samples (nasopharyngeal swab). At the top, the samples corresponding to symptomatic individuals show a bimodal distribution which is characteristic of COVID-19 [29,45]. A preliminary inner analysis of this data suggests that no correlation exists between Ct value (value that estimates the viral load) and symptomatic severity, according to other findings [46] although a high correlation was described for IgG and IgM serological tests. On the bottom panel, the asymptomatic samples show a similar distribution, also bimodal. The average Ct is indeed slightly lower in the asymptomatic cases, although no significant differences are apparent. The log dependence of Ct values obscures the fact that the viral loads comprise a span of 4 million-fold among both symptomatic and asymptomatic individuals. As a rule of thumb, derived from different tests based in culture infectivity, Ct below 30 are considered infectious. Both distributions have two thirds of the positive samples in this zone, while the very high infective, superspreader-like viral loads of Ct<20 (1000-fold higher or more) constitute the 20% of the symptomatic individuals and more than 30% of the asymptomatic ones.
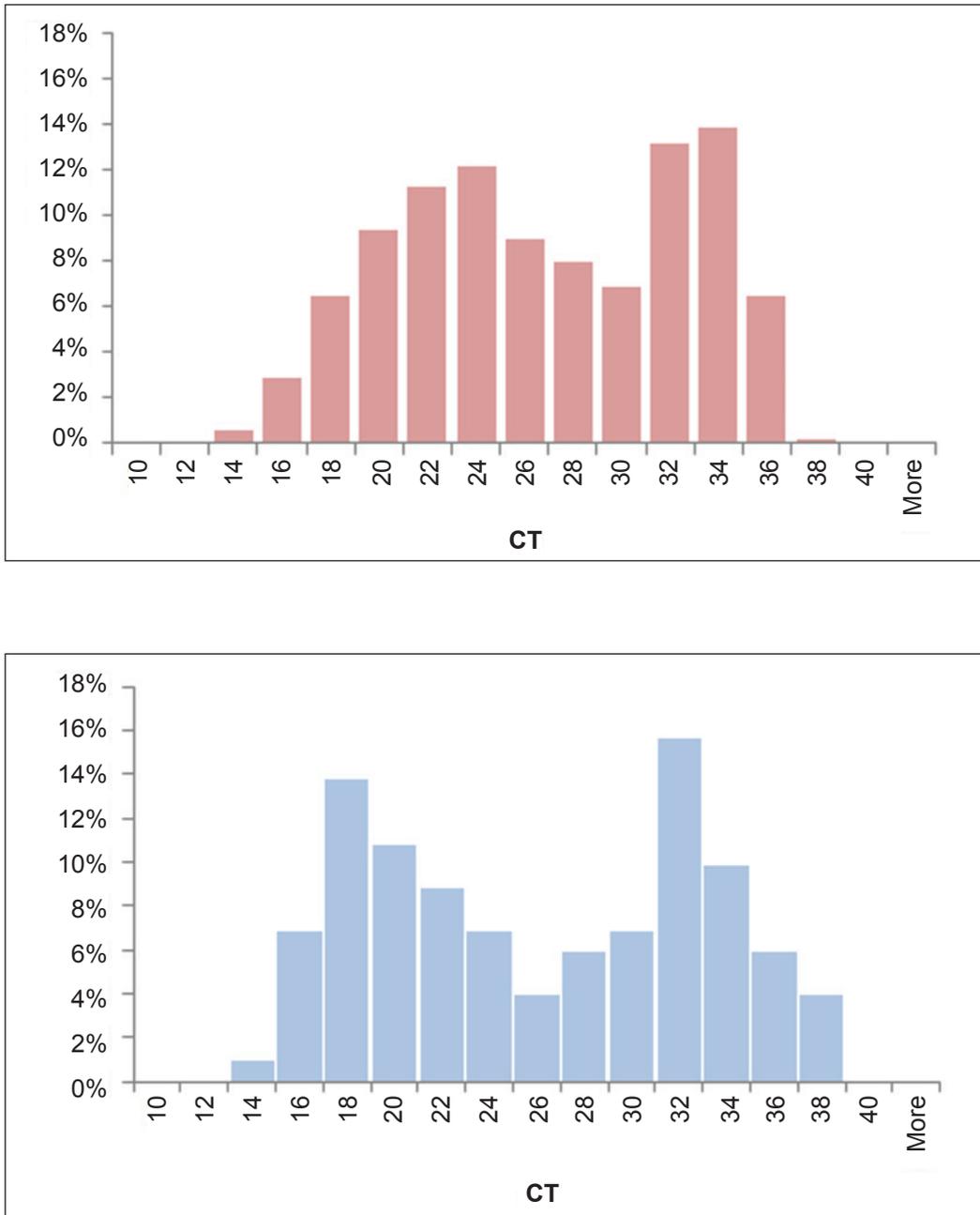
Figure 5. Distribution of cycle threshold (Ct) for N gene in Top: symptomatic individuals (N=526) and Bottom: asymptomatic individuals (N = 102).

Given this insight, and the now recognized fact that asymptomatic manifestation of COVID-19 constitutes the main fraction of all cases [47], it immediately follows that active surveillance of asymptomatic infected individuals becomes mandatory in order to be effective in the task of killing the contagion dynamics. For a sensitive and very selective, but rather slow and expensive test, as RT-qPCR is, pool strategy can be a good choice, especially in low prevalence conditions.

From the comparison of the detection probability and the complete histogram of Ct values obtained from the clinical samples, we estimated the robustness of the pool testing method. In this sense, the probability of positive detection of a single sample in a 20-samples pool was assessed. The recovery value was 95.3%. The 5% that was lost was not evenly distributed among the samples but corresponded to the lower viral loads of the specimens [29]. The very low Limit of Detection (LoD) of this technique allows even very high numbers of individual samples per pool. Together with a fast reaction from the health system, pooled testing of asymptomatic individuals can be the more cost-effective choice.

Concluding remarks

The implementation of tests in pooled samples for COVID-19 allow the expansion of detection capabilities in the community, as well as in groups of close contacts, such as hospital institutions, or institutions for the care of people in general. This methodology not only saves consumables but also time, which is a key aspect for making timely decisions of health systems. Our experience has shown that the use of pooling strategy in an active surveillance allows for the early detection COVID-19 outbreaks, increasing the likelihood of saving lives especially in places where risk groups are concentrated. As the capacity of the health systems for follow-up the pandemic is limited because of its magnitude, trials in most countries generally focus on patients with severe, or at least obvious symptoms, while potentially infectious carriers in the community remain undetected with the consequent spread of the virus. The detection of SARS-CoV-2 infected individuals including the asymptomatic/presymptomatic or oligosymptomatic (APO) can be approached from the laboratory with strategies such as those described here. As expected, the pooling strategy works best in settings with low prevalence and cohesive individual subgroups with low clinical suspicion such as APO individuals. The use of this strategy is not only of crucial importance during the pandemic in the absence of vaccination against COVID-19, but also upon vaccination in order to be able to monitor the infection particularly, in vaccinated populations who may not undergo the disease but do present infection.

Acknowledgements

References

[1] Sudre, C.H., Lee, K.A., Lochlainn, M.N., Varsavsky, T., et al., Symptom clusters in COVID-19: A potential clinical prediction tool from the COVID Symptom Study app. *Sci. Adv.* 2021, 7, eabd4177.

[2] Yang, J., Zhao, X., Liu, X., Sun, W., et al., Clinical Characteristics and Eosinophils in Young SARS-CoV-2-Positive Chinese Travelers Returning to Shanghai. *Front. Public Heal.* 2020, 8.

[3] Gao, M., Yang, L., Chen, X., Deng, Y., et al., A study on infectivity of asymptomatic SARS-CoV-2 carriers. *Respir. Med.* 2020, 169.

[4] Li, H., Wang, Y.Y., Ji, M., Pei, F., et al., Transmission Routes Analysis of SARS-CoV-2: A Systematic Review and Case Report. *Front. Cell Dev. Biol.* 2020, 8, 618.

[5] Zhou, R., Li, F., Chen, F., Liu, H., et al., Viral dynamics in asymptomatic patients with COVID-19. *Int. J. Infect. Dis.* 2020, 96, 288–290.

[6] Lavezzo, E., Franchin, E., Ciavarella, C., Cuomo-Dannenburg, G., et al., Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature* 2020, 584, 425–429.

[7] Gudbjartsson, D.F., Helgason, A., Jonsson, H., Magnusson, O.T., et al., Spread of SARS-CoV-2 in the Icelandic Population. *N. Engl. J. Med.* 2020, 382, 2302–2315.

[8] García, Z., Taylor, L., Ruano, A., Pavón, L., et al., Evaluation of a pooling method for routine anti-HCV screening of blood donors to lower the cost burden on blood banks in countries under development. *J. Med. Virol.* 1996, 49, 218–222.

[9] Cardoso, M.S., Koerner, K., Kubanek, B., Mini-pool screening by nucleic acid testing for hepatitis B virus, hepatitis C virus, and HIV: Preliminary results. *Transfusion* 1998, 38, 905–907.

[10] Kacena, K.A., Quinn, S.B., Howell, M.R., Madico, G.E., et al., Pooling urine samples for ligase chain reaction screening for genital Chlamydia trachomatis infection in asymptomatic women. *J. Clin. Microbiol.* 1998, 36, 481–485.

[11] Kacena, K.A., Quinn, S.B., Hartman, S.C., Quinn, T.C., Gaydos, C.A., Pooling of urine samples for screening for Neisseria gonorrhoeae by ligase chain reaction: Accuracy and application. *J. Clin. Microbiol.* 1998, 36, 3624–3628.

[12] Lefrère, J.J., Coste, J., Defer, C., Mercier, B., et al., Screening blood donations for viral genomes: Multicenter study of real- time simulation using pooled samples on the model of hepatitis C virus RNA detection. *Transfusion* 1998, 38, 915–923.

[13] Morandi, P.A., Schockmel, G.A., Yerly, S., Burgisser, P., et al., Detection of human immunodeficiency virus type 1 (HIV-1) RNA in pools of sera negative for antibodies to HIV-1 and HIV-2. *J. Clin. Microbiol.* 1998, 36, 1534–1538.

[14] Roth, W.K., Weber, M., Seifried, E., Feasibility and efficacy of routine PCR screening of blood donations for hepatitis C virus, hepatitis B virus, and HIV-1 in a blood-bank setting. *Lancet* 1999, 353, 359–363.

[15] Yerly, S., Pedrocchi, M., Perrin, L., The use of polymerase chain reaction in plasma pools for the concomitant detection of hepatitis C virus and HIV type 1 RNA. *Transfusion* 38, 908–914.

[16] Chong, B.S.W., Tran, T., Druce, J., Ballard, S.A., et al., Sample pooling is a viable strategy for SARS-CoV-2 detection in low-prevalence settings. *Pathology* 2020, 52, 796–800.

[17] Cao, S., Gan, Y., Wang, C., Bachmann, M., et al., Post-lockdown SARS-CoV-2 nucleic acid screening in nearly ten million residents of Wuhan, China. *Nat. Commun.* 2020, 11, 5917.

[18] Freire-Paspuel, B., Vega-Mariño, P., Velez, A., Cruz, M., Garcia-Bereguiain, M.A., "Sample pooling of RNA extracts to speed up SARS-CoV-2 diagnosis using CDC FDA EUA RT-qPCR kit." *Virus Res.* 2020, 290, 198173.

[19] Lohse, S., Pfuhl, T., Berkó-Göttel, B., Rissland, J., et al., Pooling of samples for testing for SARS-CoV-2 in asymptomatic people. *Lancet Infect. Dis.* 2020, 20, 1231–1232.

[20] Nyazika, T.K., Kaela, R., Mugoni, M., Musomekwa, K., et al., Implementation of Antibody Rapid Diagnostic Testing versus Real-Time Reverse Transcription-PCR Sample Pooling in the Screening of COVID-19: a Case of Different Testing Strategies in Africa. *mSphere* 2020, 5.

[21] Gupta, E., Padhi, A., Khodare, A., Agarwal, R., et al., Pooled RNA sample reverse transcriptase real time PCR assay for SARS CoV-2 infection: A reliable, faster and economical method. *PLoS One* 2020, 15, e0236859.

[22] Yelin, I., Aharony, N., Tamar, E.S., Argoetti, A., et al., Evaluation of COVID-19 RT-qPCR Test in Multi sample Pools. *Clin. Infect. Dis.* 2020, 71, 2073–2078.

[23] Lim, K.L., Johari, N.A., Wong, S.T., Khaw, L.T., et al., A novel strategy for community screening of SARS-CoV-2 (COVID-19): Sample pooling method. *PLoS One* 2020, 15.

[24] Abid, S., Ferjani, S., El Moussi, A., Ferjani, A., et al., Assessment of sample pooling for SARS-CoV-2 molecular testing for screening of asymptomatic persons in Tunisia. *Diagn. Microbiol. Infect. Dis.* 2020, 98, 115125.

[25] Taylor, L., Uruguay is winning against covid-19. This is how. *BMJ* 2020, 370.

[26] Hogan, C.A., Sahoo, M.K., Pinsky, B.A., Sample Pooling as a Strategy to Detect Community Transmission of SARS-CoV-2. *JAMA* 2020, 323, 1967–1969.

[27] Abdalhamid, B., Bilder, C.R., McCutchen, E.L., Hinrichs, S.H., et al., Assessment of Specimen Pooling to Conserve SARS CoV-2 Testing Resources. *Am. J. Clin. Pathol.* 2020, 153, 715–718.

[28] Ben-Ami, R., Klochendler, A., Seidel, M., Sido, T., et al., Large-scale implementation of pooled RNA extraction and RT-PCR for SARS-CoV-2 detection. *Clin. Microbiol. Infect.* 2020, 26, 1248–1253.

[29] Ambrosis, N., Martin Aispuro, P., Belhart, K., Bottero, D., et al., Active Surveillance of Asymptomatic, Presymptomatic, and Oligosymptomatic SARS-CoV-2-Infected Individuals in Communities Inhabiting Closed or Semi-closed Institutions. *Front. Med.* 2021, 8.

[30] Dorfman, R., The Detection of Defective Members of Large Populations. *Ann. Math. Stat.* 1943, 14, 436–440.

[31] Wolf, J.K., Born Again Group Testing: Multiaccess Communications. *IEEE Trans. Inf. Theory* 1985, 31, 185–191.

[32] Du, D.-Z., Hwang, F.K., *Pooling Designs and Nonadaptive Group Testing*, vol. Volume 18, WORLD SCIENTIFIC, 2006.

[33] Aldridge, M., Johnson, O., Scarlett, J., Group testing: An information theory perspective. *Found. Trends Commun. Inf. Theory* 2019, 15, 196–392.

[34] Armendáriz, I., Ferrari, P.A., Fraiman, D., Dawson, S.P., Group testing with nested pools. *arXiv* 2020.

[35] Eberhardt, J.N., Breuckmann, N.P., Eberhardt, C.S., Multi-Stage Group Testing Improves Efficiency of Large-Scale COVID-19 Screening. *J. Clin. Virol.* 2020, 128, 104382.

[36] Mutesa, L., Ndishimye, P., Butera, Y., Souopgui, J., et al., A pooled testing strategy for identifying SARS-CoV-2 at low prevalence. *Nature* 2020.

[37] Ghosh, S., Rajwade, A., Krishna, S., Gopalkrishnan, N., et al., Tapestry: A Single-Round Smart Pooling Technique for COVID-19 Testing. *medRxiv* 2020, 2020.04.23.20077727.

[38] Shental, N., Levy, S., Wuvshet, V., Skorniakov, S., et al., Efficient high-throughput SARS-CoV-2 testing to detect asymptomatic carriers. *Sci. Adv.* 2020, 6, 5961–5972.

[39] de Salazar, A., Aguilera, A., Trastoy, R., Fuentes, A., et al., Sample pooling for SARS-CoV-2 RT-PCR screening. *Clin. Microbiol. Infect.* 2020, 26, 1687.e1-1687.e5.

[40] Sunjaya, A.F., Sunjaya, A.P., Pooled Testing for Expanding COVID-19 Mass Surveillance. *Disaster Med. Public Health Prep.* 2020, 14, e42–e43.

[41] Garg, J., Singh, V., Pandey, P., Verma, A., et al., Evaluation of sample pooling for diagnosis of COVID-19 by real time-PCR: A resource-saving combat strategy. *J. Med. Virol.* 2020, n/a.

[42] Regen, F., Eren, N., Heuser, I., Hellmann-Regen, J., A simple approach to optimum pool size for pooled SARS-CoV-2 testing. *Int. J. Infect. Dis.* 2020, 100, 324–326.

[43] Lu, J., Gu, J., Gu, J., Li, K., et al., COVID-19 Outbreak Associated with Air Conditioning in Restaurant, Guangzhou, China, 2020. *Emerg. Infect. Dis.* 2020, 26, 1628–1631.

[44] Milton, D.K., A Rosetta Stone for Understanding Infectious Drops and Aerosols. *J. Pediatric Infect. Dis. Soc.* 2020, 9, 413–415.

[45] Barak, N., Ben-Ami, R., Sido, T., Perri, A., et al., Lessons from applied large-scale pooling of 133,816 SARS-CoV-2 RT-PCR tests. *Sci. Transl. Med.* 2021, eabf2823.

[46] Lee, S., Kim, T., Lee, E., Lee, C., et al., Clinical Course and Molecular Viral Shedding among Asymptomatic and Symptomatic Patients with SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea. *JAMA Intern. Med.* 2020, 180, 1447–1452.

[47] Day, M., Covid-19: four fifths of cases are asymptomatic, China figures indicate. *BMJ* 2020, 369, m1375.

Bios

Felipe Marceca

Obtained his degree (2016) and his PhD (2021) in Mathematics from the School of Exact and Natural Sciences at the University of Buenos Aires (FCEN-UBA), Argentina. Currently, he has a postdoctoral position at the University of Vienna, Austria, where he is conducting research in the fields of functional and harmonic analysis.

Luciana Rocha-Viegas

Obtained her degree in Biology (2000) and her PhD in Biological Chemistry (2005) from the School of Exact and Natural Sciences at the University of Buenos Aires (FCEN-UBA), Argentina. Currently, she is Associate Investigator of CONICET at the Laboratory of Nuclear Receptors and Chromatin at the Institute of Physiology, Molecular Biology and Neurosciences (IFIBYNE), CONICET and teaching assistant of molecular and cellular biology at FCEN-UBA.

Nicolás Pregi

Obtained his degree in Biology (2001, University of Buenos Aires), and has a PhD from the University of Buenos Aires (2007). Nowadays, he works as a researcher at CONICET in the field of neuroepigenetics of stress at the IQuiBiCEN (CONICET, Argentina), and as a professor at the National University of San Martín (Buenos Aires, Argentina).

María Gabriela Barbás

Obtained her degree in Biochemistry (1999) and her specialization in virology (2003) from the Chemistry school at the National University of Córdoba (FCQ-UNC), Argentina. She is currently, secretary of prevention and health promotion at the Health Ministry of Córdoba, Argentina. Also she was Head of the Central Laboratory of Córdoba, Argentina (since 2007 until 2019); teacher at the chair of microbiology of the Catholic University of Córdoba (since 2014 until 2019); Management consultant of AIDS and STD of the National Health Ministry of Argentina (since 2014 until 2019).

Daniela Hozbor

Obtained her degree in Biochemistry (1987) and her PhD in Biochemical Sciences (1993) from La Plata National University (UNLP), Argentina. Currently, she is Senior Investigator of CONICET, Head of the Laboratory of Vaccines (VacSal) at the Institute of Biotechnology and Molecular Biology (IBBM), UNLP-CONICET and Full Professor of the Biotechnology and Molecular Biology Area of the Faculty of Sciences. UNLP. She is in charge of the Pertussis Reference Laboratory and is a member of the steering committee of the Global Pertussis Initiative.

Adali Pecci

Obtained her degree in Chemistry (1988) and her PhD in Biological Chemistry (1997) from the School of Exact and Natural Sciences at the University of Buenos Aires (FCEN-UBA), Argentina. Currently, she is Senior Investigator of CONICET, Head of the Laboratory of Nuclear Receptors and Chromatin at the Institute of Physiology, Molecular Biology and Neurosciences (IFIBYNE), CONICET and Associated Professor at Biological Chemistry Department, FCEN-UBA. She is Vice-Director of the IFIBYNE and Academic Officer of FCEN-UBA.

Roberto Etchenique

Obtained his degree in Chemistry (1994) and his PhD in Inorganic, Analytical and Physical Chemistry (1998) from the School of Exact and Natural Sciences at the University of Buenos Aires (FCEN-UBA), Argentina. Currently, he is Senior Investigator of CONICET, Head of the Molecular Devices Lab at the Institute of Physical Chemistry of Materials, Environment and Energy (INQUIMAE), CONICET and Full Professor at the Inorganic, Analytical and Physical Chemistry Department, FCEN-UBA.