scientific reports

OPEN

Evaluation of efficiency and sensitivity of 1D and 2D sample pooling strategies for SARS-CoV-2 RT-qPCR screening purposes

Jasper Verwilt[1,2,3], Jan Hellemans[4], Tom Sante[2,3], Pieter Mestdagh[1,2,3,4] & Jo Vandesompele [1,2,3,4]✉

To increase the throughput, lower the cost, and save scarce test reagents, laboratories can pool patient samples before SARS-CoV-2 RT-qPCR testing. While different sample pooling methods have been proposed and effectively implemented in some laboratories, no systematic and large-scale evaluations exist using real-life quantitative data gathered throughout the different epidemiological stages. Here, we use anonymous data from 9673 positive cases to model, simulate and compare 1D and 2D pooling strategies. We show that the optimal choice of pooling method and pool size is an intricate decision with a testing population-dependent efficiency-sensitivity trade-off and present an online tool to provide the reader with custom real-time 1D pooling strategy recommendations.

One of the key strategies in the global battle against the COVID-19 pandemic is massive population testing. However, an ongoing shortage of time, reagents and testing capacity has tempered these efforts. Pooled testing of samples presents itself as a valid strategy to overcome these hurdles and to realize rapid, large-scale testing at lower cost and lower dependence on test reagents. By now, 1D pooling strategies have been widely applied for SARS-CoV-2 infection surveillance[1–3].

Multiple recent studies discussed pooling strategies in the frame of SARS-CoV-2 testing. Researchers have explored many strategies, but two of them have been welcomed for their simplicity and effectiveness: one-time pooling (1D pooling) and two-dimensional pooling (2D pooling). In 1D pooling, the samples are pooled, pools are tested and samples in positive pools are tested individually (Fig. 1)[4–7]. Labs worldwide have extensively evaluated 1D pooling strategies for SARS-CoV-2 testing in the lab[8–11] or using simulations[4]. In 2D pooling, samples are organized in a 2D matrix and pools are created along the matrix's rows and columns. The pools are tested, and negative rows and columns are excluded from the matrix. Next, all remaining samples are tested individually (Fig. 1)[12]. Other more complex strategies exist, such as repeated pooling[4], P-BEST[13] and Tapestry[14].

While attractive, pooling strategies come with inherent limitations. First, pooling dilutes each sample, possibly to such an extent that the viral RNA becomes undetectable, which results in false negative observations[11,12,15]. A second limitation is that an increase in sample manipulations augments the risk of cross-contamination and sample mix-ups, possibly leading to false negatives and false positives[12]. Last, when pooling, identifying individual positive samples will take an additional RNA extraction and RT-qPCR run, while one run is sufficient when testing individual samples without pooling, which increases labor and turnaround time.

Although the number of preprints and peer-reviewed publications on pooling strategies for COVID-19 RT-qPCR-based testing has accelerated rapidly throughout the pandemic, some critical aspects remain mostly ignored. First of all, the proposed optimal pooling strategy is most often based on a binary classification of samples as either positive or negative. This Boolean approach is not true to the real-world situation and does not investigate the pooling step's dilution effect. Second, when using Cq values as a semi-quantitative measure[16] of the viral loads, their overall distribution should reflect the real-life population. A high fraction of Cq values close to the limit of detection of the RT-qPCR assay produces an elevated risk of resulting in false negative samples[17]. Last, since the Cq distribution of the sample population and prevalence may vary over time, it remains unclear how the pooling strategy's performance evolves as the pandemic progresses.

[1]OncoRNALab, Cancer Research Institute Ghent, Corneel Heymanslaan 10, 9000 Ghent, Belgium. [2]Department of Biomolecular Medicine, Ghent University, Corneel Heymanslaan 10, 9000 Ghent, Belgium. [3]Center for Medical Genetics, Ghent University, Corneel Heymanslaan 10, 9000 Ghent, Belgium. [4]Biogazelle, Technologiepark 82, 9052 Zwijnaarde, Belgium. ✉email: jo.vandesompele@ugent.be
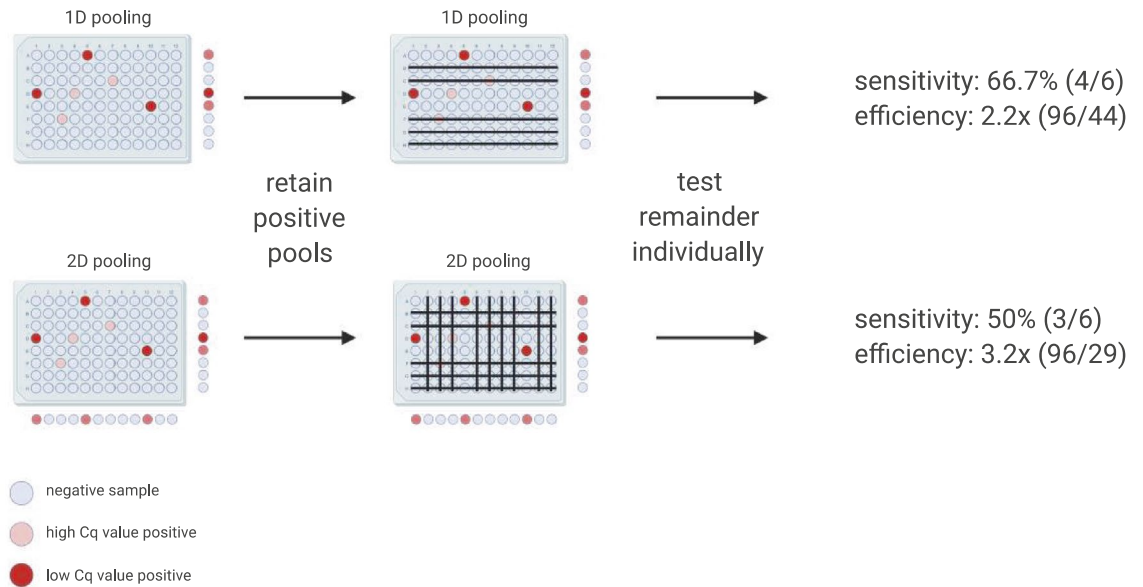
Figure 1. Schematic overview of the applied pooling strategies. The samples are represented as wells in a 96-well microtiter plate. The color of the wells indicates the samples' SARS-CoV-2 RNA concentration. In 1D pooling, the pools are created by row, the pools are tested and the samples in positive pools are tested again individually. During 2D pooling, the pools are created by row and column (each sample exists in two pools), the pools are tested, all negative rows and columns are removed and the remaining samples are tested individually. The sensitivity and the efficiency are calculated according to the equations found in the methods.

We questioned to what extent optimal pooling strategies would have changed throughout the COVID-19 pandemic and how testing facilities might use pooling strategies for future testing in a correct and attainable manner. To this extent, we simulated and evaluated one-dimensional (1D) and two-dimensional (2D) pooling strategies with different pool sizes using real-life RT-qPCR data gathered by the Belgian national testing platform during the end of the first and the beginning of the second SARS-CoV-2 epidemiological waves.

Materials and methods

Patient samples. Nasopharyngeal swabs were collected in VTM or DNA/RNA Shield (Zymo Research) by a healthcare professional as a diagnostic test for SARS-CoV-2, as part of the Belgian national testing platform. The individuals were tested at nursing homes or in triage centers, between April 9th and June 7th, and between September 1st and November 10th. After filtering the data as described further, this resulted in 207,944 patients in total, of which 9673 positives (4.65%).

SARS-CoV-2 RT-qPCR test. During the first (spring) wave, RNA extraction was performed using the Total RNA Purification Kit (Norgen Biotek #24300) according to the manufacturer's instructions using 200 μl transport medium, 200 μl lysis buffer and 200 μl ethanol, with processing using a centrifuge (5810R with rotor A-4-81, both from Eppendorf). RNA was eluted from the plates using 50 μl elution buffer (nuclease-free water), resulting in approximately 45 μl eluate. RNA extractions were simultaneously performed for 94 patient samples and 2 negative controls (nuclease-free water). After addition of the lysis buffer, 4 μl of a proprietary 700 nucleotides spike-in control RNA (prior to May 25th, 40 000 copies for singleplex RT-qPCR; from May 25th onwards, 5000 copies for duplex RT-qPCR) and carrier RNA (200 ng of yeast tRNA, Roche #10109517001) was added to all 96 wells from the plate. To the eluate of one of the negative control wells, 7500 RNA copies of positive control RNA (Synthetic SARS-CoV-2 RNA Control 2, Twist Biosciences #102024) were added. During the second (autumn) wave, RNA extraction was performed using the Quick-RNA Viral 96 Kit (Zymo Research #R1041), according to the manufacturer's instructions using 100 μl transport medium, with processing using a centrifuge (5810R with rotor A-4-81, both from Eppendorf). RNA was eluted from the plates using 30 μl elution buffer (nuclease-free water). RNA extractions were simultaneously performed for 92 patient samples, 2 negative controls (nuclease-free water), and 2 positive controls (1 diluted positive case as a full workflow control; 1 positive control RNA as RT-qPCR control, see further). After addition of the lysis buffer, 4 μl of a proprietary 700 nucleotides spike-in control RNA (5000 copies) and carrier RNA (200 ng of yeast tRNA, Roche #10109517001) was added to all 96 wells from the plate. To the eluate of one of the negative control wells, 7500 RNA copies of positive control RNA (Synthetic SARS-CoV-2 RNA Control 2, Twist Biosciences #102024) were added.

Six μl of RNA eluate was used as input for a 20 μl RT-qPCR reaction in a CFX384 qPCR instrument using 10 μl iTaq one-step RT-qPCR mastermix (Bio-Rad #1725141) according to the manufacturer's instructions, using 250 nM final concentration of primers and 400 nM of hydrolysis probe. Primers and probes were synthesized by Integrated DNA Technologies using clean-room GMP production. For detection of the SARS-CoV-2 virus, the Charité E gene assay was used (FAM)[18]; for the internal control, a proprietary hydrolysis probe assay (HEX)

was used. Prior to May 25th, 2 singleplex assays were performed; from May 25th onwards, 1 duplex RT-qPCR was performed. Cq values were generated using the FastFinder software v3.300.5 (UgenTec). Only batches were approved with a clean negative control and a positive control in the expected range.

Digital PCR calibration of positive control RNA. Digital PCR was done on a QX200 instrument (Bio-Rad) using the One-Step RT-ddPCR Advanced Kit for Probes (Bio-Rad #1864022) according to the manufacturer's instructions. Briefly, 22 µl pre-reactions were prepared consisting of 5 µl 4× supermix, 2 µl reverse transcriptase, 6 µl positive control RNA (125 RNA copies/µl), 15 mM dithiothreitol, 900 nM of each forward and reverse primer and 250 nM E gene hydrolysis probe (FAM) (see higher). 20 µl of the pre-reaction was used for droplet generation using the QX200 Droplet Generator, followed by careful transfer to a 96-well PCR plate for thermocycling: 60 min 46 °C reverse transcription, 10 min 95 °C enzyme activation, 40 cycles of 30 s denaturation at 95 °C and 1 min annealing/extension at 59 °C, and finally 10 min 98 °C enzyme deactivation. Droplets were analyzed by the QX200 Droplet Reader and QuantaSoft software. With an RNA input of 7500 copies per reaction, the digital PCR result was 1500 cDNA copies (or 20% of the expected number, a fraction confirmed by Dr. Jim Huggett for particular lot numbers of #102024, personal communication. The reason for this discrepancy is two-fold: the number of RNA molecules provided by the manufacturer is only approximate, and the reverse transcription reaction is inefficient). The median Cq value of the positive control RNA of 24.55 thus corresponds to 1500 digital PCR calibrated cDNA molecules.

Determination of efficiency and sensitivity for simulated of 1D and 2D pooling strategies. Simulations are run using R 4.0.1[19]. First, several cohorts of 100,000 patients are repeatedly simulated with varying fractions of positive cases f, depending on the fraction of positive samples of the investigated week. This is done five times, resulting in five replicate cohorts per week. The Cq values of the positive samples are sampled with replacement from the set of the positive Cq values of said week. Next, the patients are randomly separated into pools depending on the pooling strategy that is simulated. The pooling strategies that were simulated are 1×4, 1×8, 1×12, 1×16, 1×24 (all 1D), and 8×12, 12×16 and 16×24. The Cq value of the pool was calculated as follows:

$$c_{pool} = \log_2 P - \log_2 \sum_{i=1}^{p} 2^{-c_i} \tag{1}$$

With c_{pool} the Cq value of the pool, P the number of samples in the pool, p the number of positive samples in the pool, $c_1, c_2, \ldots, c_p$ the Cq values of the positive samples. If the Cq value of the pool is lower than the single-molecule Cq value, it is classified as a positive pool. For 1D pooling, only samples in positive pools are retained and the remaining individual Cq values were checked to be positive. For 2D pooling, the Cq values of the differently sized pools are checked simultaneously and the samples in negative pools are removed, after which all Cq values of the remaining samples are checked individually. Samples that were retained after the testing of the pools and that had an individual Cq lower than the single-molecule Cq value are classified as positive, all other samples are classified as negative.

The sensitivity is calculated as:

$$sensitivity = \frac{no.\ true\ positive\ samples}{no.\ true\ positive\ samples + no.\ false\ negative\ samples} \tag{2}$$

The analytical efficiency gain is calculated as:

$$efficiency\ gain = \frac{no.\ tests\ required\ for\ individual\ testing}{no.\ tests\ required\ for\ pooling\ strategy} \tag{3}$$

In all simulations, the number of tests required for individual testing is equal to the number of samples (assuming no technical failures). The outcomes for each simulation were identical as the sample size far outreached the size of the dataset. The code is available at https://github.com/OncoRNALab/covidpooling.

Ad hoc sensitivity and efficiency calculation. To calculate the efficiency for a specific 1D pooling strategy on a real sample set, the following equation was used:

$$efficiency = \frac{n}{\frac{n}{s} \cdot \left(1 + s \cdot \sum_{k=1}^{s} \left(\frac{s!}{k!(s-k)!} \cdot p^k \cdot (1-p)^{s-k} \cdot (1-c^k)\right)\right)} \tag{4}$$

With sample size n, pool size s, fraction of positive samples p and fraction of Cq values of positive samples above the 'dilution detection limit': the lowest individual Cq value that can result in a pooled Cq value lower than the single molecule Cq value, or:

$$single\ molecule\ Cq\ value - \log_2(pool\ size) \tag{5}$$

Equation (4) is derived as follows. The efficiency is defined by the following equation:

$$efficiency = \frac{n}{no.\ tests\ required\ for\ pooling\ strategy} \tag{6}$$

The number of tests performed when using a pooling strategy is equal to:

$$no.\ tests\ required\ for\ pooling\ strategy = no.\ pools + no.\ positive\ pools \cdot s \tag{7}$$

Since $\#\ pools = \frac{n}{s}$,

$$no.\ tests\ required\ for\ pooling\ strategy = \frac{n}{s} + no.\ positive\ pools \cdot s \tag{8}$$

The exact number of positive pools can be calculated by multiplying the number of pools by the probability of a pool testing positive. Approximately, a pool will test positive if it includes a positive sample with a Cq value lower than the 'dilution detection limit'. The probability of having a specific number of positive samples k in a pool with pool size s is defined by a binomial distribution:

$$\frac{s!}{k!(s-k)!} \cdot p^k \cdot (1-p)^{s-k} \tag{9}$$

Thus, the probability of having at least one positive value in a pool is equal to:

$$\sum_{k=1}^{s} \left(\frac{s!}{k!(s-k)!} \cdot p^k \cdot (1-p)^{s-k} \right) \tag{10}$$

In general, we can assume that when a sample has a Cq value higher than the 'dilution detection limit', for the sample to test positive, it must be accompanied by a sample with a Cq value lower than the 'dilution detection limit'. Equation (10) can be adjusted to factor for these events:

$$\sum_{k=1}^{s} \left(\frac{s!}{k!(s-k)!} \cdot p^k \cdot (1-p)^{s-k} \cdot \left(1-c^k\right) \right) \tag{11}$$

Filling in Eq. (11) in Eq. (9) results in the final formula being used for the calculation of the efficiency.

To estimate the sensitivity for a specific 1D pooling strategy on a real sample set, the following equation was used:

$$sensitivity = c \cdot \sum_{k=0}^{s-1} \left(\frac{(s-1)!}{k!((s-1)-k)!} \cdot p^k \cdot (1-p)^{(s-1)-k} \cdot \left(1-c^k\right) \right) + (1-c) \tag{12}$$

The sensitivity can be defined as the probability a true positive sample tests positive. For our situation it will be equal to the probability that any sample tests positive:

$$P(pos\ test) = P(pos\ test|Cq \geq cut\ off) \cdot P(Cq \geq cut\ off) + P(pos\ test|Cq < cut\ off) \cdot P(Cq < cut\ off) \tag{13}$$

Previously, $P(Cq \geq cut\ off)$ was defined as c and therefore $P(Cq < cut\ off) = 1 - c$. Also $P(pos\ test|Cq < cut\ off) = 1$. A positive sample with Cq value above the 'dilution detection limit' can only test positive if one of the other samples in the pool is also positive and has a Cq value lower than the 'dilution detection limit'. We can calculate the probability of this happening by using the same logic as before, but with $s-1$ instead of s:

$$\sum_{k=0}^{s-1} \left(\frac{(s-1)!}{k!((s-1)-k)!} \cdot p^k \cdot (1-p)^{(s-1)-k} \cdot \left(1-c^k\right) \right) \tag{14}$$

Completing Eq. (13) with Eq. (14) leads to Eq. (12) for calculating the sensitivity.

Web application. To help laboratories find the best pooling strategy for their specific situation (i.e. the local positivity ratio and Cq value distribution), we developed a Shiny application in R 4.0.1. The Shiny application was launched on our in-house Shiny server and is available at https://shiny.dev.cmgg.be/.

Ethical declarations. Our analyses have been approved under EC/071-2020 by the Ghent University Hospital ethical committee. No biological samples were used in this study and the analyses were based solely on anonymized data gathered by the Belgian national testing platform. The need for an informed consent was waived by the ethics committee.

Results

Single-molecule Cq value determination. To infer the single-molecule Cq value (specific to our laboratory set-up), we made a 5-point tenfold serial dilution series of positive control RNA from 150,000 (digital PCR calibrated) copies down to 15 copies. The Y-intercept value of the linear regression points at a single-molecule Cq value of 35.66 and 35.28 for singleplex and duplex RT-qPCR (which are the two methods that were used, see "Materials and methods"), respectively (Supplemental Fig. 1). Therefore, we conservatively use 37 as the
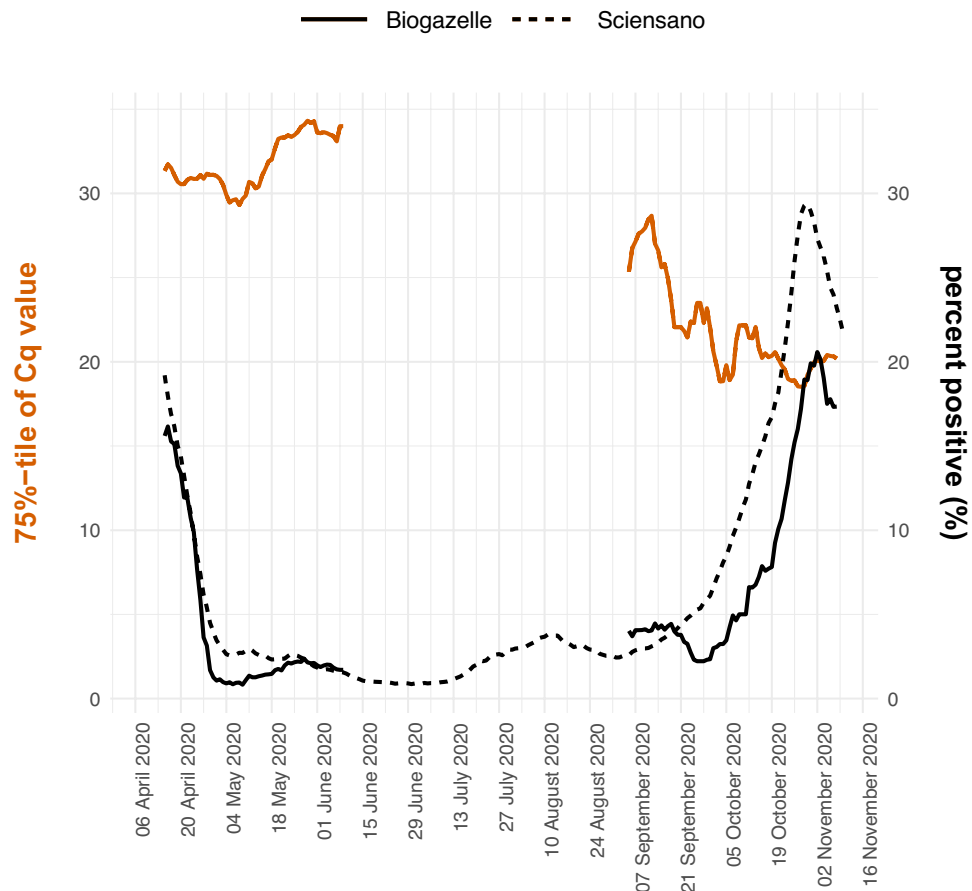
Figure 2. Evolution of the 75%-tile of the Cq value distribution and fraction of positive samples. The left y-axis shows the 7 day moving window average of the 75%-tile of the Cq value distribution of the data originating from Biogazelle and the right y-axis shows the 7 day moving window average of the fraction of positive samples for the Biogazelle and Sciensano data. The two datasets are differentiated by the line type. If the moving average was calculated using on the basis of less than 5 days (due to no data being available for specific days), the datapoint was removed from the visualization.

single-molecule value for further analysis. Patient sample Cq values higher than the single-molecule Cq value threshold are likely due to random measurement variation, lot reagent variability and sample inhibition. Note that this cut off is highly dependent on the experimental procedure and reagents used and is, therefore, specific to our experimental setting.

Cq distribution is dynamic over course of the pandemic. Few studies have explored how the Cq value distribution within one testing facility evolves during the COVID-19 pandemic. We determined the 75%-tile of the Cq value distribution (i.e. 75% of the data have a lower Cq value) and the percentage of positive tests per day as a proxy for actual Cq value distribution and prevalence, respectively (Fig. 2). We compared the fraction of positive tests in our dataset with the fraction of positive tests as reported by the federal agency for public health Sciensano (https://epistat.wiv-isp.be/covid/. accessed January 25th, 2021). First, the fractions of positive tests seem to align at the end of the first wave, but in the second wave our data seems to be shifted about 1–2 weeks later. Second, the 75%-tile of the Cq values varies over the course of the pandemic from a minimum value of around 18 and a maximum value of almost 35. Third, when comparing the fraction of positive samples and the 75%-tile of the Cq value distribution, we note that these parameters are inversely related: when the positivity rate goes down, the Cq value distribution shifts towards the higher end of the spectrum (Supplemental Fig. 2). In conclusion, the Cq value distribution and prevalence show a dynamic profile over the course of the COVID-19 pandemic. These observations are crucial considering that positivity rate and Cq value distribution are key determinants of efficiency and sensitivity of any pooling strategy.

Pooling efficiency and sensitivity changes as pandemic progresses. To explore how hypothetical pooling strategies would have affected the SARS-CoV-2 testing outcomes, we simulated different 1D (with pool size of 4, 8, 12, 16 and 24) and 2D pooling (with pool sizes of 8 × 12, 12 × 16, and 16 × 24) strategies using individual sample Cq values from a single Belgian laboratory during the end of the first and beginning of the second
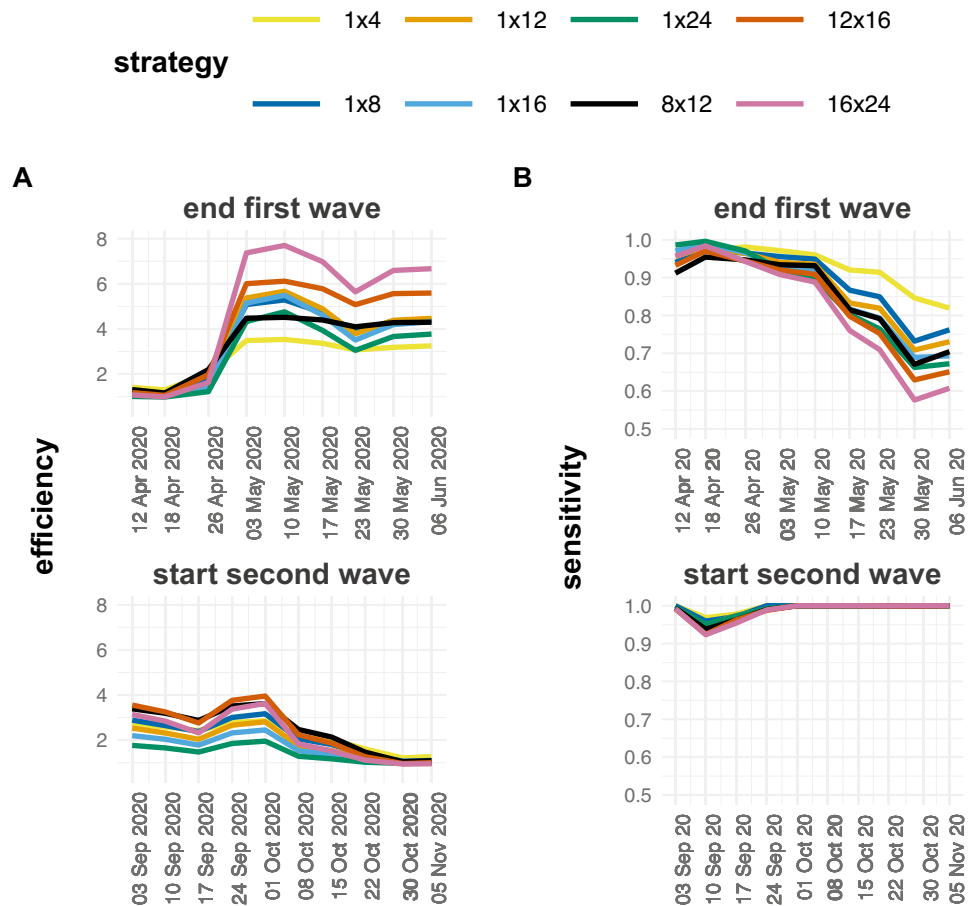
Figure 3. Sensitivity and efficiency for the end of the first (**A**) and the start of the second (**B**) Belgian SARS-CoV-2 infection wave. The data is grouped by week and the sensitivity and efficiency are calculated by simulating different pooling strategies (1×4, 1×8, 1×12, 1×16, 1×24, 8×12, 12×16 and 16×24). The pooling strategies can be distinguished by color.

wave. The data were grouped by week and the resulting Cq value distributions and positivity rates were used as input for the simulations (Fig. 3). First, sensitivity and efficiency show very opposing patterns when comparing different timeframes during the pandemic. At the end of the first wave the efficiency increases, while at the beginning of the second wave, the efficiency decreases. The sensitivity drops as we move further away from the first wave but remains stable as we enter the second. Second, pool size and strategy have a major influence on the outcomes. 2D pooling strategies generally have the highest efficiency, but the lowest sensitivity. Curiously, strategies with larger pool sizes were more efficient during the end of the first wave, but less efficient during the beginning of the second wave. The sensitivity was always higher for strategies with smaller pool sizes, irrespective of the time during the pandemic. We conclude that—just like the positivity rate and the Cq value distribution—the sensitivity and efficiency depend on the timing in the pandemic and are heavily affected by the pooling strategy and the size of the pools.

Positivity rate drives efficiency, Cq distribution drives sensitivity. We wondered how the positivity rate, Cq value distribution and pooling strategy affect the performance of the adopted strategy. To investigate this, we used the previous simulations for the end of the first wave to create an adjusted visualization where all parameters involved are incorporated (Fig. 4). First, it is apparent that weeks with a high 75%-tile Cq value tend to have a low sensitivity and weeks with a high positivity rate seem to have a low efficiency. Second, pooling strategies with smaller pool sizes seem less sensitive to changes in positivity rate and Cq value distribution, as indicated by the area of the polygon traced around the edges of the data (Fig. 4). These results show that the prevalence mainly contributes to the efficiency and the Cq distribution to the sensitivity.

Shiny app for guided decision making. To provide laboratories with a custom pooling strategy recommendation based on their specific sampling population, we worked out equations to estimate the sensitivity and efficiency (for 1D pooling strategies) based on an uploaded dataset of Cq values. The derivation of these equations can be found in "Materials and methods" section. We focused on 1D pooling strategies since 2D pooling strategies generally resulted in extreme outcomes (highest efficiency and lowest sensitivity) and the outcomes of the optimal pooling strategy are situated somewhere in two extremes. Thus, we provide no tool for 2D pooling
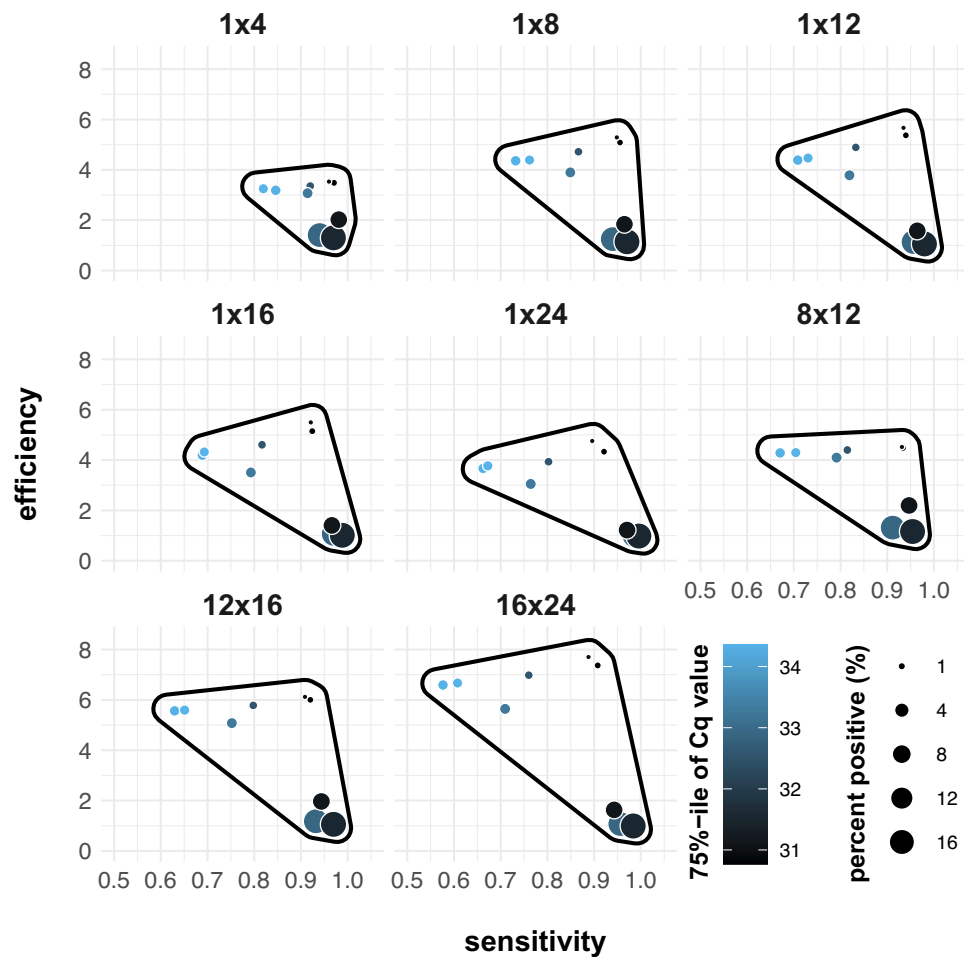
Figure 4. Simulated sensitivity and efficiency for the end of the first wave visualized with relation to the week (different circles), fraction of positive samples (size of circles) and 75%-tile of the Cq value distribution (color). A polygon is drawn around the datapoints (with a small margin) to visualize and to compare the variability of the sensitivity and efficiency over a period of time between pooling strategies.

methods. To evaluate the equations' capacities to replicate the simulations, we compared the simulated efficiency and sensitivity of the pooling strategies for the different weeks and the efficiency and sensitivity of the pooling strategies the distributions, fraction of positive samples and single-molecule cut off as inputs for the formulas (Supplemental Figs. 3, 4). We integrated these formulas into an open-access Shiny application (Supplemental Fig. 5). The application provides a step-by-step guide to completing the input variables and parameters. A Frequently Asked Questions (FAQ) section answers some questions on the workings and general idea of the application. The tool requires four inputs: a dataset of Cq values from positive samples, the positivity rate, the single-molecule cut off Cq value and a range of pool sizes of interest. The Shiny application will then swiftly output the estimated data-specific efficiency and sensitivity for different pooling strategies. The application shows a graphic representation of the dataset-specific change in sensitivity (green) and efficiency (black) for the requested range of pool sizes. The output also indicates the pool sizes resulting in maximum efficiency and sensitivity. The tool can be easily updated if comments or suggestions arise.

Discussion

Using a sizeable real-life dataset of 9673 SARS-CoV-2 positive nasopharyngeal samples, we found that the pooling strategies' sensitivity and efficiency mainly depend on the prevalence and the distribution of the Cq values. Our results indicate that both the prevalence and the Cq value distribution are dynamic parameters during the SARS-CoV-2 pandemic and that, as a result, the resulting sensitivity and efficiency of pooling strategies are as well. To enable researchers and institutions with a real-time and accessible recommendation concerning the optimal 1D pooling strategy for their testing population, we developed a Shiny app providing just that.

Two factors could explain the dynamics of the prevalence and the Cq value distribution: epidemiological and virological change within the same sampling population and variation in the sampling population. The existence of these factors would suggest that an intricate interplay of these two components is at the origin of the observed evolutions. Recent research indicated that the first component (epidemiological change) exists, as the distribution of random surveillance testing-deduced Cq values fluctuates during the SARS-CoV-2 pandemic (by definition,

no changes in sampling population occurred in this research, thereby excluding this factor from the equation)[20]. As such, an alarming change in reproduction number (inherently coupled to the epidemiological period) should induce a reassessment of the pooling strategy[21]. The second component (variation in sampling population) is bound to happen when the testing facility is not consistently receiving samples from the same origin, as is the case for Biogazelle. At the very introduction of Biogazelle as a testing facility, most samples originated from hospitals and sources were added progressively as the testing capacity increased. Additionally, the Belgian government instituted a rapid change in the testing regime on October 21st, 2020: only symptomatic suspected SARS-CoV-2 cases get tested. The federal government lifted this measure on November 23rd, 2020, when the number of cases lowered and the existing testing capacity sufficed again. Since symptomatic patients generally show lower Cq values[22,23], it is clear that sampling bias will contribute to the overall Cq value distribution.

The influence these dynamic parameters have on the variation of performance of pooling strategies is significant. This observation raises an issue for interpreting pooling strategy evaluations not based on time-series datasets. The effectiveness of a chosen pooling plan might even decrease to such an extent that it becomes inferior to individual testing. We observed this situation at the end of the second wave when efficiency is close to 1, but sensitivity is not (Fig. 3). Based on these results, it becomes essential to regularly re-evaluate an adopted pooling strategy to avoid compromising on sensitivity and efficiency when there is no need.

Multiple effects contribute to how the testing population's characteristics drive pooling strategy outcomes. The main trends show that the prevalence mainly influences efficiency, and the Cq value distribution mainly influences sensitivity (Fig. 3). We can explain both observations by using common sense and basic mathematics. When the prevalence is low, the efficiency is high: fewer pools will have positive samples and therefore test negative, which will automatically result in a lower number of tests needed to test all samples. Additionally, when a considerable proportion of samples have a Cq value close to the single-molecule Cq value, a more significant fraction of samples will become too diluted to detect during pooling and result in false negatives. There appear to be secondary compensating effects of the Cq value distribution and prevalence on the efficiency and sensitivity, respectively, which are more subtle. Primarily, as a higher fraction of positive samples has a Cq value close to the upper limit, more pools will test (false) negative, boosting the efficiency. On the other hand, when the prevalence increases, the sensitivity will increase due to an effect we call 'rescuing': a high Cq value that would otherwise test negative when diluted in the pool is 'rescued' by a low Cq value in the same pool. When the prevalence rises, the chances of this phenomenon happening also increase and as will the sensitivity. The same was observed by Cleary et al.[21]. Although minor, these secondary effects explain several of our observations.

To elaborate how the optimal pooling strategy (best efficiency trade-off) transforms over time, assume two situations: low prevalence and high prevalence. When the prevalence is low, the larger pool sizes will result in higher efficiency and lower prevalence (more dilution). However, when the prevalence is high, the 'rescuing' effect will be more prominent and counteract the increasing efficiency and decreasing sensitivity. These results are in line with the widely accepted idea that sample pooling methods show a higher efficiency when pool size is large and that as prevalence increases, it reached a threshold after which smaller pool sizes become more efficient[4,12]. Intuitively, the 'rescuing' effect is less prominent in 2D pooling strategies, as both pools (row and column) need to rescue the high Cq sample.

False negatives have pre-pool Cq values close to the detection limit and predominantly originate from patients who are at the end of an infection[21,24], putting their clinical relevance in question (i.e. no longer infectious). Similarly, however, one can argue that these high Cq samples are imperative to a favorable pandemic response: they might originate from pre-symptomatic or very recently-infected patients[21], allowing for catching cases before transmission—a principle at the very core of every population screening strategy. Also, we cannot rule out that these high Cq values are due to imperfect sampling or any other mistakes along the sample preparation[16].

The preference of achieved efficiency and sensitivity depends on the viral circulation. When the virus is widely circulating, and many tests are conducted, efficiency will be more critical as missing several positive samples will not have a remarkable effect. On the other hand, when the number of cases is extremely low, sensitivity should be prioritized: missing one sample can initiate a break-out. Disturbingly, screening settings are a combination of both previous settings: the number of tests is high, but the positivity rate is low. Such conditions require a purposely calculated efficiency-sensitivity trade-off.

Our study suffers from some essential limitations. First, although the data grouped by weeks provides many different situations to assess, there will still be other combinations of parameters that we did not analyze in this paper. However, the current dataset probably represents the most plausible scenarios as the data originate from a protracted period of the pandemic. Second, we selected only 1D and 2D pooling methods in this simulation study. As stated before, other pooling regimes exist and might be more performant than the discussed ones. Yet, these pooling strategies come with intrinsic shortcomings. The P-BEST pooling protocol is very time consuming[13], even when using a pipetting robot, and the repeated pooling method suffers from a complicated re-pooling scheme[4]. Third, our model relies on the critical assumption that we can directly induce the pool's Cq value from the individual samples' Cq values using a simple formula (see "Materials and methods"). Our pooling results are solely based on simulations (while the individual Cq values are experimentally determined) and wet lab experiments have shown that these do not necessarily correspond perfectly[8–11]. Fourth, to calculate the pooling strategies' performance, the single-molecule Cq value and the prevalence must be known. However, we can easily calculate the single-molecule Cq value by generating a dilution series of a calibrated reference material as done in this study. Of note, this should be done in each lab individually as it depends on the experimental set-up. A testing laboratory can also choose to utilize a cut off Cq value different from the single-molecule Cq value. Once a threshold Cq value is determined, it should not be changed. The prevalence, however, cannot be known precisely, and as a result, the prevalence must be estimated. We can do this either before adopting a pooling strategy by testing the individual samples and using the fraction of positive samples as an indication for the prevalence or when a pooling strategy is already in place by calculating it from the percentage of positive pools[5,21]. Last, the

calculated efficiency gain is merely a representation of the number of individual RNA extractions and RT-qPCR reactions and does not evaluate the amount of labor or time-to-result. Pooling a low number of samples will unnecessarily increase the time-to-result and workload.

In conclusion, we show that finding the optimal pooling strategy for SARS-CoV-2 test samples is guided by a testing population-dependent efficiency-sensitivity trade-off. Consequently, the most favorable pooling regime might change throughout the pandemic due to epidemiological changes and revisions in diagnostic testing strategies. We provide an accessible shiny application to guide readers towards the optimal pooling strategy to fit their needs.

Data availability

The code and Cq value data are available on https://github.com/OncoRNALab/covidpooling.

References

1. Dewald, F. *et al.* Effective high-throughput RT-qPCR screening for SARS-CoV-2 infections in children. *MedRxiv* https://doi.org/10.1101/2022.02.04.22270304 (2022).
2. Barak, N. *et al.* Lessons from applied large-scale pooling of 133,816 SARS-CoV-2 RT-PCR tests. *Sci. Transl. Med.* **13**, eabf2823 (2021).
3. Pollock, N. R. *et al.* Implementation of SARS-CoV2 screening in K–12 schools using in-school pooled molecular testing and deconvolution by rapid antigen test. *J. Clin. Microbiol.* **59**, 25 (2021).
4. Shani-Narkiss, H., Gilday, O. D., Yayon, N. & Landau, I. D. Efficient and practical sample pooling high-throughput PCR diagnosis of COVID-19. *medRxiv* https://doi.org/10.1101/2020.04.06.20052159 (2020).
5. Guha, P., Guha, A. & Bandyopadhyay, T. Application of pooled testing in screening and estimating the prevalence of Covid-19. *MedRxiv* https://doi.org/10.1101/2020.05.26.20113696 (2020).
6. Adams, K. Expanding covid-19 testing: Mathematical guidelines for the optimal sample pool size given positive test rate. *medRxiv* https://doi.org/10.1101/2020.05.21.20108522 (2020).
7. Millioni, R. & Mortarino, C. Sequential informed pooling approach to detect SARS-CoV2 infection. *medRxiv* https://doi.org/10.1101/2020.04.24.20077966 (2020).
8. Hogan, C. A., Sahoo, M. K. & Pinsky, B. A. Sample pooling as a strategy to detect community transmission of SARS-CoV-2. *J. Am. Med. Assoc.* **323**, 1967–1969 (2020).
9. Yelin, I. *et al.* Evaluation of COVID-19 RT-qPCR test in multi-sample pools. *Clin. Infect. Dis.* https://doi.org/10.1093/cid/ciaa531 (2020).
10. Abdalhamid, B. *et al.* Assessment of specimen pooling to conserve SARS CoV-2 testing resources. *Am. J. Clin. Pathol.* **153**, 715–718 (2020).
11. Torres, I., Albert, E. & Navarro, D. Pooling of nasopharyngeal swab specimens for SARS-CoV-2 detection by RT-PCR. *J. Med. Virol.* **92**, 2306–2307 (2020).
12. Sinnott-Armstrong, N., Klein, D. & Hickey, B. Evaluation of group testing for SARS-CoV-2 RNA. *medRxiv* https://doi.org/10.1101/2020.03.27.20043968 (2020).
13. Shental, N. *et al.* Efficient high-throughput SARS-CoV-2 testing to detect asymptomatic carriers. *Sci. Adv.* **6**, 5961–5972 (2020).
14. Ghosh, S. *et al.* A compressed sensing approach to group-testing for COVID-19 detection. *arXiv* (2020).
15. Gan, Y. *et al.* Sample pooling as a strategy of SARS-COV-2 nucleic acid screening increases the false-negative rate. *medRxiv* https://doi.org/10.1101/2020.05.18.20106138 (2020).
16. Dahdouh, E., Lázaro-Perona, F., Romero-Gómez, M. P., Mingorance, J. & García-Rodriguez, J. Ct values from SARS-CoV-2 diagnostic PCR assays should not be used as direct estimates of viral load. *J. Infect.* https://doi.org/10.1016/j.jinf.2020.10.017 (2020).
17. Buchan, B. W. *et al.* Distribution of SARS-CoV-2 PCR cycle threshold values provide practical insight into overall and target-Specific sensitivity among symptomatic patients. *Am. J. Clin. Pathol.* https://doi.org/10.1093/AJCP/AQAA133 (2020).
18. Corman, V. M. *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* **25**, 2000045 (2020).
19. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2021).
20. Hay, J. A., Kennedy-Shaffer, L., Kanjilal, S., Lipsitch, M. & Mina, M. J. Estimating epidemiologic dynamics from single cross-sectional viral load distributions. *medRxiv* https://doi.org/10.1101/2020.10.08.20204222 (2020).
21. Cleary, B. *et al.* Using viral load and epidemic dynamics to optimize pooled testing in resource-constrained settings. *Sci. Transl. Med.* https://doi.org/10.1126/scitranslmed.abf1568 (2021).
22. Singanayagam, A. *et al.* Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Eurosurveillance* https://doi.org/10.2807/1560-7917.ES.2020.25.32.2001483 (2020).
23. Gorzalski, A. J. *et al.* Characteristics of viral specimens collected from asymptomatic and fatal cases of COVID-19. *J. Biomed. Res.* **34**, 431–436 (2020).
24. Tom, M. R. & Mina, M. J. To interpret the SARS-CoV-2 test, consider the cycle threshold value. *Clin. Infect. Dis.* **71**, 2252–2254 (2020).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-10581-6.

Correspondence and requests for materials should be addressed to J.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.