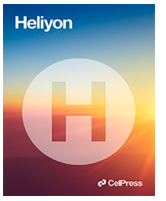
Heliyon 11 (2025) e41623

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Reagent efficiency and analytical sensitivity optimization for a reliable SARS-CoV-2 pool-based testing strategy

José P. Miranda [a,b,c], Javiera Osorio [a], Marcia Silva [a], Carola Silva [a], Victoria Madrid [a], Rossana Camponovo [a], Marcela Henríquez-Henríquez [a,b,d,*]

[a] Bupa Lab, part of Bupa, La Florida, Santiago, Chile
[b] Advanced Center for Chronic Diseases (ACCDiS), Pontificia Universidad Católica de Chile & Universidad de Chile, Santiago, Chile
[c] Department of Nutrition, Diabetes, and Metabolism, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
[d] Millennium Nucleus in Cardiovascular Magnetic Resonance, Millennium Science Initiative Program, ANID, Santiago, Chile

ARTICLE INFO

Keywords:
SARS-CoV-2
Pool testing
Efficiency
Sensitivity
Public health

ABSTRACT

Background: The SARS-CoV-2 pandemic caused millions of infections worldwide. Among the strategies for effective containment, frequent and massive testing was fundamental. Although sample pooling allows multiplying the installed analysis capacity, the definition of the number of samples to include in a pool is commonly guided more by economic parameters than analytical quality.

Methods: We developed a mathematical model to determine the pooling conditions that maximize reagent efficiency and analytical sensitivity. We evaluated 30 samples individually and in 2-sample to 12-sample pools. Using Passing Bablok regressions, we estimated the shift of Ct values in the RT-qPCR reaction for each pool size. With this Ct shift, we estimated sensitivity in the context of the distribution of 1,030 individually evaluated positive samples.

Findings: Our results showed that the most significant gain in efficiency occurred in the 4-sample pool, while at pools greater than 8-sample, there was no considerable reagent savings. Sensitivity significantly dropped to 87.18 %–92.52 % for a 4-sample pool and reached as low as 77.09 %–80.87 % in a 12-sample pooling.

Conclusions: Our results suggest that a 4-sample pooling maximizes reagent efficiency and analytical sensitivity. These considerations are essential to increase testing capacity and efficiently detect and contain contagious.

1. Introduction

The COVID-19 pandemic posed a constant challenge to public health globally. One of the most critical challenges was maintaining the diagnosis, whose gold standard was the amplification of the genetic material of the SARS-CoV-2, using the Reverse Transcription Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) technique [1]. Due to the worldwide shortage of reagents required for the multiple steps of this technique [2], modifications arose, such as direct protocols [3–7] that escape the step of viral RNA extraction and the analysis of samples in pools [8–13]. This late modification involved analyzing a discrete number of samples from different patients in the same reaction, before or after RNA extraction, followed by RT-qPCR. When a pool gives a positive result, each of the

* Corresponding author. Bupa Lab, part of Bupa, Chile.
E-mail address: marcela.henriquezh@bupa.cl (M. Henríquez-Henríquez).

https://doi.org/10.1016/j.heliyon.2025.e41623
Received 5 October 2023; Received in revised form 30 November 2024; Accepted 1 January 2025
Available online 2 January 2025

samples that compose it is reprocessed individually (2-sample hierarchical pooling).

Pooling strategies made it possible to multiply the installed PCR capacity in many laboratories. However, an immanent risk of pooling is to guide pool configuration decisions to maximize economic efficiency without considering the impact of pooling on analytical quality.

This work aimed to establish conditions that simultaneously maximize economic efficiency with analytical sensitivity, considering that massive testing and high-quality RT-qPCR diagnostic or screening are essential for effectively managing a pandemic [14].

2. Materials and methods

2.1. Efficiency estimation

Given a number of samples (*N)* to be analyzed in laboratory routine under a positivity proportion (*P)*, the expected number of positive samples (*EPS*) can be estimated by:

$$EPS = N\ x\ P \tag{1}$$

Defining pooling size (*PS*) as the number of samples analyzed together in a single reaction, and assuming the simplest scenario in which every positive sample will be distributed in a different pool, the number of samples to repeat (*STR*) in each analysis can be estimated according to:

$$STR = EPS\ x\ PS \tag{2}$$

Then, the total number of samples to process per analysis (*STP*) is determined by:

$$STP = N + STR \tag{3}$$

The minimum expected reagent efficiency (*E*), that is, the minimal fraction of total sample processing reactions that are saved using a specific pooling size, can be expressed as:

$$E = 1 - \frac{STP}{N\ x\ PS} \tag{4}$$

Replacing with the previous terms, the minimum expected reagent efficiency can be expressed as a function of positivity proportion (*P*) and the pooling size (*PS*):

$$E = 1 - P - \frac{1}{PS} \tag{5}$$

To maximize the efficiency function, the second derivative of equation (5) can be drawn as follow:

$$E'' = -\frac{2}{PS^3} \tag{6}$$

2.2. Clinical specimens

To test the effect of different pool configurations (pooling size) on sensitivity and the shift on Ct values from individual and pooled samples, we aleatory selected 30 positive samples. To define the historical distribution of Ct's, we analyzed 3,392 samples from the first peak of infections in Chile (June–July 2020), of which 1,030 yielded a positive result, and 2,362 were negative (30.4 % positivity). All samples were obtained from the clinical routine at Bupa Lab, part of Bupa, Santiago - Chile. Samples were collected using naso-pharyngeal swabs and stored at 4 °C in tubes containing sterile Weise's buffer, as described before [7], until analysis. This project was approved by the Ethics Committee of IntegraMédica, part of Bupa, Chile (number BL01-102023). All procedures followed the Helsinki Declaration. All samples were anonymized, and patients signed an informed consent to approve using their anonymized results for epidemiological vigilance and research.

2.3. Pool formation, nucleic acid extraction, and RT-qPCR

To form pools, we equivolumetrically mixed each positive sample with a mixture of confirmed negative samples until reaching a dilution of 2X, 4X, 6X, 8X, 10X, and 12X. We extracted the RNA of the SARS-CoV-2 from each pool using the MagNA Pure 96 DNA and Viral NA SV Kit (Roche Diagnostics, Cat. No. 06374891001) in the MagNA Pure 96 System (Roche Diagnostics).

For RT-qPCR reactions, we used the SARS-CoV-2/SARS-CoV Multiplex REAL-TIME PCR Detection Kit (DNA-Technology, Cat. No. R3-P436-23/9EU R3-P436-S3/9EU), according to the manufacturer's instructions, and loaded in a 96-well DTprime thermal cycler (DNA Technology). The DNA Technology kit allows the simultaneous amplification of 3 targets (SARS-CoV-like, SARS-CoV-2 gene E, and SARS-CoV-2 gene N) with a maximum theoretical cycle threshold (Ct) of 45. Samples were considered positive when there was amplification in all three targets.

We used the same extraction and amplification protocol for the positive samples analyzed individually as a reference for pooling, and those samples were used to evaluate the historical distribution of Ct's.

2.4. Estimation of sensitivity loss

We used the shift of Ct's observed when the same positive samples were analyzed individually versus pooled conditions, to estimate the loss of sensitivity as a function of different pooling sizes. With this in mind, we first used Passing Bablok regressions [14] and compared the Ct's from the original samples to those obtained according to each pooling size. This also allowed to estimate a theoretical limit of detection for the different pool sizes using the following equation:

$$Ct\ at\ limit\ [pool] = \alpha + \beta\ x\ Ct\ at\ limit\ [original\ sample] \quad (7)$$

Where α and β corresponded to the intercept and slope, respectively, estimated with the Passing Bablok regressions, and the Ct at the limit of detection for the original sample was empirically defined as the highest Ct obtained in our laboratory routine for a positive sample (Ct 41.6 for CoV-Like, 42.1 for CoV N, and 41.5 for CoV E genes, respectively).

Then, we characterized the historical distribution of Ct's of the 1,030 positive samples. Based on the estimated Ct shift at the limit of detection and the Ct distribution of the positive retrospective samples, we calculated the proportion of samples that would be lost (false negative) by using the pool-based RT-qPCR strategy.

All the descriptive statistics, statistical analyzes, and graphs were performed using R.

3. Results

3.1. Reagent efficiency estimation

We first obtained equation (5) to estimate how the RT-qPCR reaction efficiency behaves in a pool-based strategy at different pooling sizes and a variable test positivity. This equation is modeled in Fig. 1 (left), considering a pool size between 2 and 12 samples and positivity between 0.01% and 30%. As expected, as the degree of pooling increases or the positivity decreases, there is an increase in reagent efficiency. However, at a positivity below 1%, there is no more significant increase in efficiency. This information is summarized in Table 1. On the one hand, at a positivity of 50% and 2-sample pooling, an efficiency of 0% is expected; that is, there is no additional benefit when using a pooling strategy under these theoretical epidemiological conditions. Under positivity conditions below 1% and a pooling size of 10 or 12 samples, the efficiency exceeds 90%.

To maximize the efficiency function, we obtained the second derivative of equation (5), giving rise to equation (6). This equation, modeled in Fig. 1 (right, black line), shows that the most significant efficiency gain occurs in pooling sizes from 4 to 8 samples. At pooling sizes greater than 8, there is no real greater saving of RT-qPCR reactions.

3.2. Estimation of sensitivity loss

Using the DNA Technology multiplex RT-qPCR kit, we analyzed 30 known positive samples from the clinical routine at different pooling sizes (2 to 12). The results are described for SARS-CoV-2 N, E, and CoV-Like genes in Supplementary Tables 1, 2, and 3, respectively.

As expected, we obtained a progressive increase in Ct as the pooling size increased, with a systematic loss of sensitivity in those samples where the original Ct was closer to the detection limit (Ct > 34). When comparing the median Ct of these 30 samples between

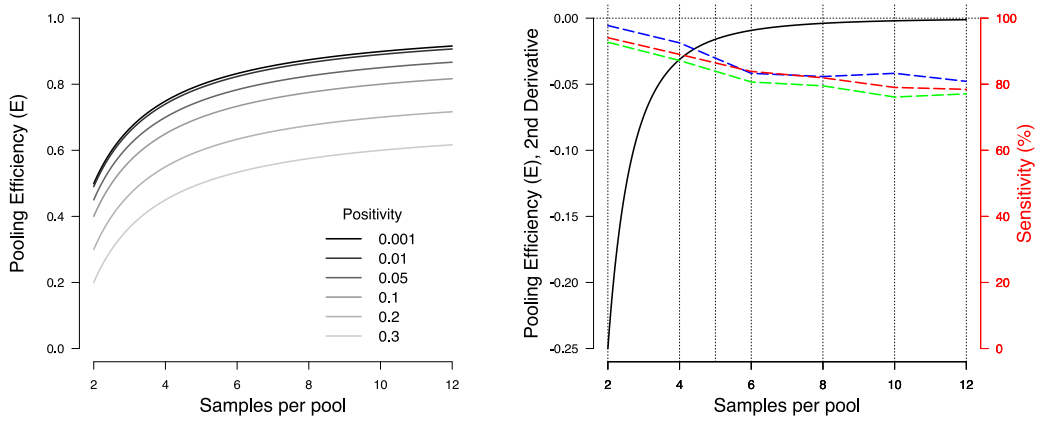


Fig. 1. Estimated reagent efficiency and sensitivity of a pool-based RT-qPCR strategy. The theoretical proportion of reagent efficiency according to the pool size at a given positivity is modeled on the left, considering equation (5). On the right, there is the intercept between the maximization of efficiency (equation (6), black line) and the estimated sensitivity at different pooling sizes on 1,030 positive retrospective samples for genes SARS-CoV-2 E (dashed red line), SARS-CoV-Like (dashed green line), and SARS-CoV-2 N (dashed blue line).

Table 1
Reagent efficiency at different scenarios of positivity and pooling size. RT-qPCR reaction savings were estimated using equation (5).

Positivity (%)	Expected positive samples/1000 samples	Efficiency (E) per pooling size					
		2X	4X	6X	8X	10X	12X
50	500	0	0.25	0.33	0.38	0.40	0.42
30	300	0.20	0.45	0.53	0.58	0.60	0.62
10	100	0.40	0.65	0.73	0.78	0.80	0.82
5	50	0.45	0.70	0.78	0.83	0.85	0.87
1	10	0.49	0.74	0.82	0.87	0.89	0.91
0.1	1	0.50	0.75	0.83	0.87	0.90	0.92

each pool size and the original sample, we found a shift ranging from +1.4 to +1.8 cycles for the 2-sample pool, which increases from +4.0 to +4.3 cycles for the 12-sample pool.

Considering that the obtained Ct does not necessarily represent the laboratory routine distribution, a sensitivity estimated from this low-sample analysis may be subject to significant variations. A more robust strategy to estimate sensitivity lies in the shift of the Ct's regarding a historical distribution; therefore, it could be used as a more realistic approximation to assess the sensitivity loss according to the pooling size.

Although the supplier indicates that the DNA Technology kit's detection limit is Ct 45, in practice, the highest value in our laboratory routine was Ct 42.1, 41.5, and 41.6 for SARS-CoV-2 N, E, and CoV-Like genes, respectively. Given that we know that each sample should have yielded a positive result in Supplementary Tables 1, 2, and 3, for the summary statistics regarding negative samples, we assigned those values as our practical limit of detection.

Therefore, we estimated the Ct shift using Passing Bablok regressions for each pooling size concerning the original sample using the Ct's of Supplementary Tables 1, 2, and 3 (Supplementary Figs. 1, 2, and 3). In general, we found a high correlation with the original sample (Pearson's $\rho > 0.989$), with narrow confidence intervals for the Passing Bablok regressions that widen slightly at the lower and upper Ct values. We found that for a 2X pooling size, there is a shift ranging from +1.46 to +1.62 cycles, with an estimated loss of sensitivity from Ct 39.98 to 40.64 (Table 2). Considering the historical distribution of Ct's for the 1,030 positive samples, the percentage of samples that could be missed at this estimated detection limit is 2.23%, 5.92%, and 7.28%; therefore, the estimated sensitivity for this degree of pooling was 97.77%, 94.08%, and 92.72% for SARS-CoV-2 N, E, and CoV-Like genes, respectively. Given that the Passing Bablok regression's confidence intervals for 2X pool size include the value 0 in the intercept and the value 1 in the slope, this loss of sensitivity is not statistically different from the original sample (Table 2).

For pool sizes 4X to 12X, we found a shift ranging between +2.2-2.6 and +4.0–4.3 cycles, respectively. Therefore, the estimated Ct where sensitivity is lost was between 39.15-39.70 and 37.31–37.73, respectively (Table 2). Again, if we consider the historical distribution of Ct's in 1,030 positive samples, the estimated percentage of false-negative samples resulting from a 4X pooling strategy would be 7.48%, 10.97%, and 12.82%, and a sensitivity of 92.52%, 89.03%, and 87.18% for SARS-CoV-2 N, E, and CoV-Like genes, respectively. Regarding the 12X pool size, the sensitivity would be reduced to 80.87%, 78.45%, and 77.09%, respectively (Table 2). Since none of the confidence intervals for the intercepts of the Passing Bablok regressions in the pool sizes from 4X to 12X contains the value 0, the difference in the loss of sensitivity is statistically significant concerning the original sample.

The Ct's distribution in the 1,030 positive retrospective samples previously confirmed in the laboratory routine is shown in Fig. 2. The Ct's have a bimodal distribution, with the first peak around Ct 21 and a second peak reaching Ct 38, which apparently would not be associated with age, sex, or COVID-19 symptoms (data not shown). This result was similar for the three genes amplified with the DNA-Technology kit, and it shows that a relevant fraction of the samples (>40%) could at risk of being false-negative through the use of a pooling strategy (Ct > 34).

3.3. Reagent efficiency versus sensitivity

When we superimposed the estimated values of sensitivity with those of reagent use efficiency, we found that a 4-sample pool size captured the highest efficiency with the lowest loss of sensitivity; therefore, this is the most reliable and efficient condition for working with pooled samples, obtaining a low rate of false-negative, and maximizing the reagent savings (Fig. 1 right, dashed red, blue, and green lines).

4. Discussion

Implementing a SARS-CoV-2 analysis in pooled samples is a strategy that considerably increases the testing capacity for diagnostic and screening purposes [15]. Many studies have currently described ways to maximize the efficiency of the pooling approach; however, there needs to be more in describing when that efficiency can retain a high sensitivity to ensure that the strategy has an acceptable quality. In the present study, we modeled different positivity and pool size scenarios and determined that a 4-sample pool size retained the highest sensitivity and reagent efficiency.

Table 2
Estimated sensitivity at different pooling sizes. Estimation was performed considering the Ct shifting from the Passing Bablok regressions between the original sample (No pool), and the distribution of 1,030 positive retrospective samples.

Gene	Pool Size	Passing Bablok intercept (95%CI)	Passing Bablok slope (95%CI)	ΔCt from original sample (equation (7))	Estimated Ct at the loss of sensitivity	Estimated sample lost from the historical distribution	Estimated Sensitivity
CoV-2 N	No pool	reference	reference	reference	42.1	0%	100%
	2X	0.556 (−0.89; 1.85)	1.021 (0.97; 1.07)	1.46	40.64	2.23%	97.77%
	4X	2.400 (0.82; 3.71)	1.000 (0.95; 1.07)	2.40	39.70	7.48%	92.52%
	6X	1.774 (0.60; 2.85)	1.054 (1.02; 1.10)	4.05	38.05	16.70%	83.30%
	8X	2.410 (1.47; 3.29)	1.041 (1.00; 1.08)	4.15	37.95	17.67%	82.33%
	10X	3.331 (1.60; 4.92)	1.017 (0.96; 1.09)	4.04	38.06	16.70%	83.30%
	12X	2.914 (0.68; 5.12)	1.034 (0.94; 1.13)	4.37	37.73	19.13%	80.87%
CoV-2 E	No pool	reference	reference	reference	41.5	0%	100%
	2X	1.281 (−0.30; 2.88)	1.006 (0.95; 1.07)	1.52	39.98	5.92%	94.08%
	4X	2.866 (0.86; 4.88)	0.985 (0.91; 1.07)	2.25	39.25	10.97%	89.03%
	6X	2.721 (0.95; 4.32)	1.013 (0.96; 1.07)	3.25	38.25	16.12%	83.88%
	8X	3.550 (1.78; 5.34)	1.000 (0.92; 1.07)	3.55	37.95	18.06%	81.94%
	10X	4.100 (2.82; 6.26)	1.000 (0.92; 1.05)	4.10	37.40	20.97%	79.03%
	12X	3.687 (1.05; 6.28)	1.011 (0.91; 1.11)	4.14	37.36	21.55%	78.45%
CoV-Like	No pool	reference	reference	reference	41.6	0%	100%
	2X	0.299 (−1.02; 1.60)	1.032 (0.98; 1.09)	1.62	39.98	7.28%	92.72%
	4X	1.787 (0.02; 3.59)	1.016 (0.94; 1.09)	2.45	39.15	12.82%	87.18%
	6X	1.472 (−0.31; 3.03)	1.051 (0.99; 1.12)	3.60	38.00	19.32%	80.68%
	8X	2.228 (0.50; 3.41)	1.039 (0.99; 1.11)	3.84	37.76	20.49%	79.51%
	10X	2.694 (1.02; 4.52)	1.041 (0.96; 1.11)	4.40	37.20	23.88%	76.12%
	12X	2.688 (0.33; 5.13)	1.038 (0.95; 1.14)	4.29	37.31	22.91%	77.09%

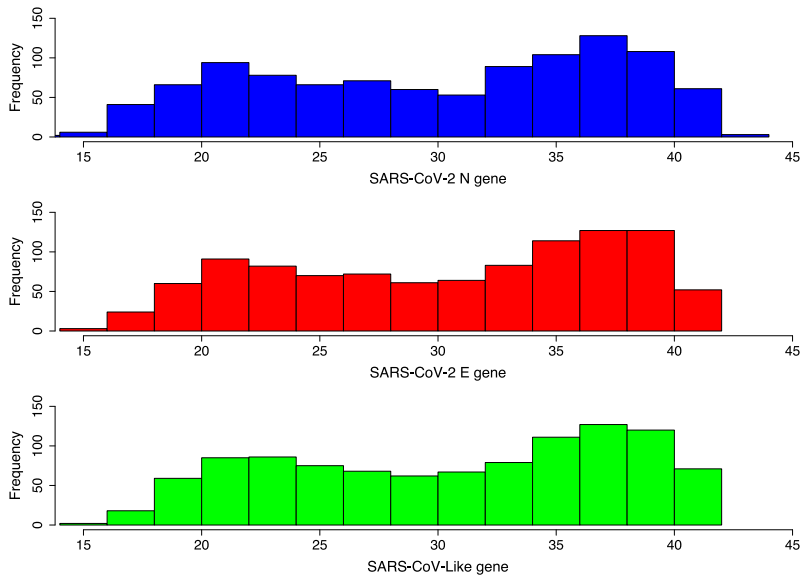


Fig. 2. Distribution of Ct's from 1,030 retrospective positive individual samples. RT-qPCR was performed using the DNA Technology kit, and results for the SARS-CoV-2 N, E, and SARS-CoV-Like genes are plotted.

4.1. Disentangling reaction efficiency models

To obtain our simplified minimal efficiency equation, that is, the minimal number of RT-qPCR reactions saved by the pooling strategy, we assumed that each positive sample was detected in a different RT-qPCR plate well. This assumption is valid when positivity is low but deviates slightly from reality, underestimating efficiency as positivity increases. In a prospective analysis in our laboratory, including 3,484 samples and a 4-sample pool strategy (871 pools), we found 190 positive pools, of which 176 (92.6%) contained only one positive sample, 13 pools (7.4%) had two positive samples, and only one pool (0.57%) had three positive samples. Sample positivity reached 5.9%, and according to our equation, the expected efficiency was 69.1%, while the practical efficiency increased slightly to 72.7% (data not shown). Generally, a 4-sample pool efficiency is greater than 25% under very high positivity conditions (50%) and increases over 75% with positivity close to or less than 0.1%.

Various equations considering positivity and the pool size have been described to determine the pooling efficiency. In several works [8,16,17], the mean efficiency estimates are based on Black et al. [18], which has open access online version (www.chrisbilder.com/shiny/). When evaluating our previous 5.9% positivity conditions with a 4-sample pool, assuming a kit sensitivity and specificity of 100%, and a 2-stage hierarchical pooling algorithm, the efficiency calculated is 53%, which considerably underestimates our practical efficiency of 72.7%. We found a similar result when considering the Dorfman efficiency equation [19], with an efficiency of 53.4% of saved reactions. Another approach to efficiency according to positivity is provided by Aragón-Caqueo et al. [20]. Using this model for the abovementioned conditions, the efficiency calculated is 46.6%, underestimating our practical efficiency also. We believe the problem underlying these estimates may be governed by the lack of integration of an important quality parameter for pool size selection: sensitivity.

4.2. The downside of direct sensitivity estimates

Sample pooling can occur at the time of sample collection (swab pooling). It can also be done in the laboratory, either upon receipt of the samples and prior to RNA extraction, or when the RNA is already purified. Each method has its pros and cons, so its selection will depend on the objective. For example, pooling nasopharyngeal swabs facilitates epidemiological surveillance in sentinel centers, since it reduces the materials required for these purposes in the field. On the other hand, when the purpose is clinical diagnosis, independent sample collection with subsequent pooling in the laboratory is more advisable, since it ensures better sample traceability and integration with computer systems for its entry and reporting. We have also reported a SARS-CoV-2 detection methodology that escapes RNA extraction, however, because this strategy already involves an increase in Ct values, we believe that the direct sample pool strategy may have a high impact on the loss of sensitivity [7].

Regardless of the pooling methodology, the sensitivity estimation due to the pooling strategy is one of the weaknesses of many studies, mainly because of the low number of samples used for their analyses [9–11], where the Ct's are not representative of the distribution of Ct's obtained in the laboratory routine. In fact, the biased use of samples with low Ct's or by mathematical models has led some researchers to affirm that the use of a pooling strategy of 20X, 30X, 50X, or even 100X is feasible [11–13,21], which is undoubtedly unthinkable for diagnostic and it may have a reduced utility in screening due to the tremendous proportion of samples considered false-negative.

Although the ideal way to estimate sensitivity for a pool is to prospectively analyze a representative number of samples, in practice this is difficult due to the high cost involved in evaluating each condition independently. Because the loss of sensitivity is almost entirely due to a systematic positive shift in high Ct's near the limit of detection, we believe that an inexpensive strategy to estimate sensitivity closer to the real world is to consider the shift in Ct's and to evaluate it in the context a retrospective distribution of individually evaluated samples. Considering this approach in a 4-sample pooling strategy compared to individual sample analysis, we found that the RT-qPCR sensitivity significantly decreased to 87.18%–92.52%. In contrast, performing a direct estimate only from the 30 positive samples gives a biased sensitivity of 100% (Supplementary Tables 1, 2, and 3). Modeling for the estimation of sensitivity based on the Ct's has also been described by Bateman et al. [22], who found a sensitivity of 93% for a pool of 5 samples; however, they mention that a weakness of their study was that dilutions were performed with VTM medium instead of confirmed negative samples. This experimental condition could have made the reaction more favorable, overestimating the sensitivity. However, the differences with our analysis are reduced and could be explained by the use of a different RT-qPCR kit or collection media.

4.3. Is positivity over 10% a limiting factor for the pooling strategy?

Regarding the limit of acceptable positivity for the pooling strategy to remain efficient, the conception that it should be 10% or less is ingrained [8,15]. However, our efficiency equation demonstrates that even with a higher positivity, saving a significant number of RT-qPCR reactions is still possible. We believe that limitations in high positivity conditions are caused more to each laboratory's technical capacity, especially concerning the delay in the turnaround time. This problem could be partially solved if the laboratory has the capacity to allocate a greater number of personnel and equipment for analysis.

4.4. Tackling the pandemic: Competent containment of contagious

The most critical factor in stopping the COVID-19 pandemic was the ability to effectively detect and contain those who could transmit the virus [23]. To this end, it is crucial to differentiate virus nucleic acid detection tests for diagnostic purposes from those for screening. According to the CDC [15], the use of a diagnostic test is justified in those people for whom there is a reasonable suspicion that the person is infected, either because of presenting symptoms or because of traceability with close contacts. On the other hand, a screening test is applied to asymptomatic or pre-symptomatic people who do not know if they were exposed to the virus. Hence, it seeks to detect infection in people who could be contagious, mainly by active case finding [24].

The RT-qPCR is a widely used diagnostic technique due to its high sensitivity and specificity; however, for screening purposes, other nucleic acid detection techniques, such as loop-mediated isothermal amplification (RT-LAMP) or the detection of antibodies have also been used. During the pandemic, implementing pooling strategies with RT-qPCR considerably reduced costs, facilitating its use for screening. Although it is believed that sensitivity could have a low impact compared to parameters such as the amount/frequency of tests and turn-around time [23] over pandemic containment, the motivation only to maximize efficiency in low-resource countries suggests that it is essential to reconsider it. Due to the urgency to increase the testing capacity, the FDA authorized diagnostic kits with a 4-sample pooling strategy; however, it is common to find that more samples are analyzed per pool for diagnostic or screening in low-resource countries.

Regarding Ct and contagiousness, the initial idea was that at Ct's over 34, a viral culture was not viable [25,26]; however, there is controversy regarding this point, considering that other studies have been able to culture the virus in symptomatic patients with Ct over 35, estimating that 8.3% of these cases could transmit the virus to other people [27]. The current view is that Ct's can vary considerably between each RT-qPCR kit, which makes it an unreliable marker of infectivity, and its values should not be used to discriminate contagious patients unless the kit is validated for this purpose [28,29].

With this information, a pooling strategy of 10 or more samples could detect most infectious patients; however, it is crucial to establish the minimum acceptable quality for its use to contain the pandemic more effectively. Our suggestion is to use a 4-sample pool since this strategy is the one that captures the highest sensitivity and efficiency, as well as detecting a significant proportion of infectious, allowing to reduce costs to test more and more frequently, in addition to achieving relatively fast delivery of results.

Sample pooling is a strategy that can significantly improve surveillance efforts for emerging and re-emerging diseases, especially in countries with limited resources or in highly exposed populations. This methodology, which is currently used for the detection of SARS-CoV-2, has also been used for other diseases. Philip et al. showed that it is possible to pool 10 to 80 serum samples for HIV detection, while Pereira et al. showed that it is also possible to pool sputum samples for Tuberculosis screening in prisons [30,31]. Both studies agree that the number of samples to be included in a pool is inversely proportional to the prevalence of the disease to be tested. This suggests that the pooling strategy should be adapted to different epidemiological conditions such as disease prevalence and pathogen evolution, as well as laboratory conditions such as sampling method, sample type, extraction of genetic material, and the test used.

4.5. Strengths and weaknesses of the study

This study integrates historical data to balance economic efficiency and sensitivity to suggest a policy for a reliable analysis using the pooling strategy. Regarding the limitations of this study, it is essential to mention the retrospective nature of estimating sensitivity loss. This approach was chosen as the primary analysis, considering that a prospective analysis for each sample and each pool size would be very costly and time-consuming to be representative of the distribution of Ct's.

In this context, it is important to consider that Ct values in retrospective studies may be affected by unmeasured confounding. For

example, the inherent variability in historical data, including different sample collection and processing protocols, as well as different RT-qPCR kits used. This variability may introduce bias and affect the accuracy of Ct values, however, since the beginning of the COVID-19 pandemic we have maintained the same analysis strategies. Despite this, there are other variables to consider that may affect the distribution of Ct values, related to the emergence of new SARS-CoV-2 variants. For example, Li et al. described that patients infected with the Delta variant had significantly lower Ct values compared to those infected with previous variants [32]. This suggests that the ability to analyze pooled samples should not be conceived as a static strategy but should be evaluated over time to fit better to the epidemiological conditions of each territory. During this study, the original SARS-CoV-2 lineage was the prevalent globally, with the Alpha variant appearing a few months later, therefore, we cannot affirm that with more virulent variants the number of samples in a pool could have been increased. Mutations in viral genome sequences can affect the binding of primers used in RT-qPCR assays. If mutations are present in the target regions of primers, this can reduce amplification efficiency and consequently increase Ct values [33]. This is further evidence of variables that can affect the pooling strategy over time, therefore, evaluation of primer efficiency according to prevalent SARS-CoV-2 variants is necessary.

It is also important to mention that the results are based on evaluating a single RT-qPCR kit; therefore, we cannot establish whether the behavior will be similar for other extraction or amplification kits. In any diagnostic testing, including for SARS-CoV-2, the reagents used for nucleic acid extraction and PCR amplification play a significant role in the accuracy, sensitivity, and overall performance of the test. This is particularly important when pooling samples, as the pooling process can dilute the viral RNA, potentially masking the detection of low viral loads.

5. Conclusion

Based on our results, we can conclude that using a 4-sample pool maximizes the efficiency and sensitivity to deliver a reliable result in detecting the SARS-CoV-2 when using a pooling strategy and represents the most favorable conditions to contain the COVID-19 pandemic.

CRediT authorship contribution statement

José P. Miranda: Writing – review & editing, Writing – original draft, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Javiera Osorio:** Supervision, Investigation, Formal analysis. **Marcia Silva:** Investigation, Formal analysis. **Carola Silva:** Investigation, Formal analysis. **Victoria Madrid:** Investigation, Formal analysis. **Rossana Camponovo:** Investigation. **Marcela Henríquez-Henríquez:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization.

Ethics statement

This study was approved by the Scientific Ethics Committee of IntegraMédica, part of Bupa, Chile on October 2, 2023 (approval number BL01-102023). The Scientific Ethics Committee approved a waiver of informed consent.

Funding

This study was funded by the Research, Development & Innovation Programme - RDI, Bupa Lab, part of Bupa, Chile.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e41623.

References

[1] CDC. Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel for Emergency Use Only, 2019. https://www.who.int/docs/default-source/coronaviruse/uscdcrt-pcr-panel-for-detection-instructions.pdf.

[2] M.N. Esbin, O.N. Whitney, S. Chong, A. Maurer, X. Darzacq, R. Tjian, Overcoming the bottleneck to widespread testing: a rapid review of nucleic acid testing approaches for COVID-19 detection, RNA 26 (7) (2020) 781–783, https://doi.org/10.1261/rna.076232.120.

[3] M.R. Hasan, F. Mirza, H. Al-Hail, S. Sundararaju, T. Xaba, M. Iqbal, et al., Detection of SARS-CoV-2 RNA by direct RT-qPCR on nasopharyngeal specimens without extraction of viral RNA, PLoS One 15 (2020).

[4] N. Lübke, T. Senff, S. Scherger, S. Hauka, M. Andrée, O. Adams, et al., Extraction-free SARS-CoV-2 detection by rapid RT-qPCR universal for all primary respiratory materials, J. Clin. Virol. 130 (2020).

[5] C. Beltrán-Pavez, L.A. Alonso-Palomares, F. Valiente-Echeverría, A. Gaggero, R. Soto-Rifo, G.P. Barriga, Accuracy of a RT-qPCR SARS-CoV-2 detection assay without prior RNA extraction, J. Virol Methods 287 (2021).

[6] E.A. Bruce, M.L. Huang, G.A. Perchetti, S. Tighe, P. Laaguiby, J.J. Hoffman, et al., Direct RT-qPCR detection of SARS-CoV-2 RNA from patient nasopharyngeal swabs without an RNA extraction step, PLoS Biol. 18 (2020).

[7] J.P. Miranda, J. Osorio, M. Videla, G. Angel, R. Camponovo, M. Henríquez-Henríquez, Analytical and clinical validation for RT-qPCR detection of SARS-CoV-2 without RNA extraction, Front. Med. 7 (2020) 567572, https://doi.org/10.3389/fmed.2020.567572.

[8] B. Abdalhamid, C.R. Bilder, E.L. McCutchen, S.H. Hinrichs, S.A. Koepsell, P.C. Iwen, Assessment of specimen pooling to conserve SARS CoV-2 testing resources, Am. J. Clin. Pathol. 153 (2020) 715–718.

[9] I. Torres, E. Albert, D. Navarro, Pooling of nasopharyngeal swab specimens for SARS-CoV-2 detection by RT-PCR, in: J Med Virol, John Wiley and Sons Inc, 2020, pp. 2306–2307.

[10] R. Ben-Ami, A. Klochendler, M. Seidel, T. Sido, O. Gurel-Gurevich, M. Yassour, et al., Large-scale implementation of pooled RNA extraction and RT-PCR for SARS-CoV-2 detection, Clin. Microbiol. Infection 26 (2020) 1248–1253.

[11] I. Yelin, N. Aharony, E.S. Tamar, A. Argoetti, E. Messer, D. Berenbaum, et al., Evaluation of COVID-19 RT-qPCR test in multi sample pools, Clin. Infect. Dis. 71 (2020) 2073–2078.

[12] S. Lohse, T. Pfuhl, B. Berkó-Göttel, J. Rissland, T. Geißler, B. Gärtner, et al., Pooling of samples for testing for SARS-CoV-2 in asymptomatic people. Lancet Infect Dis, Lancet Publishing Group, 2020.

[13] L. Mutesa, P. Ndishimye, Y. Butera, J. Souopgui, A. Uwineza, R. Rutayisire, et al., A pooled testing strategy for identifying SARS-CoV-2 at low prevalence, Nature 589 (2021) 276–280.

[14] H. Passing, W. Bablok, A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I, J. Clin. Chem. Clin. Biochem. 21 (1983) 709–720.

[15] Interim guidance for use of pooling procedures in SARS-CoV-2 diagnostic, Screening, and Surveillance Testing, 2020. https://stacks.cdc.gov/view/cdc/106237/cdc_106237_DS1.pdf.

[16] S. Abid, S. Ferjani, A. El Moussi, A. Ferjani, M. Nasr, I. Landolsi, et al., Assessment of sample pooling for SARS-CoV-2 molecular testing for screening of asymptomatic persons in Tunisia, Diagn. Microbiol. Infect. Dis. 98 (2020) 115125.

[17] B.S.W. Chong, T. Tran, J. Druce, S.A. Ballard, J.A. Simpson, M. Catton, Sample pooling is a viable strategy for SARS-CoV-2 detection in low-prevalence settings, Pathology 52 (7) (2020) 796–800.

[18] M.S. Black, C.R. Bilder, J.M. Tebbs, Optimal retesting configurations for hierarchical group testing, J R Stat Soc Ser C Appl Stat 64 (2015) 693–710.

[19] R. Dorfman, The detection of defective members of large populations, Ann. Math. Stat. 14 (1943) 436–440.

[20] D. Aragón-Caqueo, J. Fernández-Salinas, D. Laroze, Optimization of group size in pool testing strategy for SARS-CoV-2: a simple mathematical model, J. Med. Virol. 92 (2020) 1988–1994.

[21] R. Hanel, S. Thurner, Boosting test-efficiency by pooled testing for SARS-CoV-2—formula for optimal pool size, PLoS One 15 (2020).

[22] A.C. Bateman, S. Mueller, K. Guenther, P. Shult, Assessing the dilution effect of specimen pooling on the sensitivity of SARS-CoV-2 PCR tests, J. Med. Virol. 93 (3) (2020) 1568–1572.

[23] D.B. Larremore, B. Wilder, E. Lester, S. Shehata, J.M. Burke, J.A. Hay, et al., Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening, Sci. Adv. 7 (2021).

[24] Z. Li, Q. Chen, L. Feng, L. Rodewald, Y. Xia, H. Yu, et al., Active case finding with case management: the key to tackling the COVID-19 pandemic, Lancet 396 (2020) 63–70.

[25] Scola B. La, M. Le Bideau, J. Andreani, V.T. Hoang, C. Grimaldier, P. Colson, et al., Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards, Eur. J. Clin. Microbiol. Infect. Dis. 39 (2020) 1059–1061.

[26] C. Rhee, S. Kanjilal, M. Baker, M. Klompas, Duration of severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2) infectivity: when is it safe to discontinue isolation? Clin. Infect. Dis. 26 (8) (2020) 1467–1474.

[27] A. Singanayagam, M. Patel, A. Charlett, J.L. Bernal, V. Saliba, J. Ellis, et al., Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020, Euro Surveill. 25 (2020).

[28] M.J. Binnicker, Can the severe acute respiratory syndrome Coronavirus 2 Polymerase Chain reaction cycle threshold value and time from symptom onset to testing predict infectivity? Clin. Infect. Dis. 71 (2020) 2667–2668.

[29] M.P. Romero-Gómez, S. Gómez-Sebastian, E. Cendejas-Bueno, M.D. Montero-Vega, J. Mingorance, J. García-Rodríguez, Ct value is not enough to discriminate patients harbouring infective virus, J. Infect. 0 (2020).

[30] P.J. Peters, E. Westheimer, S. Cohen, L.B. Hightow-Weidman, N. Moss, B. Tsoi, et al., Screening yield of HIV Antigen/Antibody combination and pooled HIV RNA testing for Acute HIV infection in a high-prevalence population, JAMA 315 (2016) 682–690. https://jamanetwork.com/journals/jama/fullarticle/2491636.

[31] P.C.P. Dos Santos, A. Da Silva Santos, R.D. De Oliveira, B.O. Da Silva, T.R. Soares, L. Martinez, et al., Pooling sputum samples for efficient mass tuberculosis screening in prisons, Clin. Infect. Dis. 74 (2022) 2115–2121, https://doi.org/10.1093/cid/ciab847.

[32] B. Li, A. Deng, K. Li, Y. Hu, Z. Li, Y. Shi, et al., Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 Delta variant, Nat. Commun. 13 (2022) 1–9. https://www.nature.com/articles/s41467-022-28089-y.

[33] L. Cao, T. Xu, X. Liu, Y. Ji, S. Huang, H. Peng, et al., The impact of accumulated mutations in SARS-CoV-2 variants on the qPCR detection efficiency, Front. Cell. Infect. Microbiol. 12 (2022) 823306.