Effectiveness of Sample Pooling Strategies for SARS-CoV-2 Mass Screening by RT-PCR: A Scoping Review

Sangeeta Deka1  Deepjyoti Kalita1

1Department of Microbiology, All India Institute of Medical Sciences, Rishikesh, Virbhadr Road, Rishikesh, Uttarakhand, India

The ongoing COVID-19 pandemic has hugely impacted the economy of many countries, and there is an acute shortage of diagnostic resources. With the exponential increase in the number of cases and necessity to screen large number of people, there is a steep increase in the demand for diagnostic kits. Pooled-sample testing is a promising strategy to screen a large population rapidly with limited resources. The aim of this work was to compile a cohesive literature review of the effectiveness and accuracy of pooled-sample testing in the detection of SARS-CoV-2 and critically analyze its limitations. Medline, Google Scholar, Embase, and preprint servers (e.g., bioRxiv) were searched for literature on pooled testing for diagnosis of COVID-19, and out of initial 60 articles/reports, nine original articles were retained. Optimal pool size (number of samples in a pool) seemed to be dependent on factors like prevalence or rate of positivity in community. In low-prevalence localities pool size of around 30 seemed to be effective as observed by some authors. All the researchers had found significant reduction in number of tests (depending on pool size, stages, and pooling design), leading to conservation of resources. Pooling can be done with extracted RNA eluate or directly with patient’s sample before extraction. This leads to further reduction in consumables, time and manpower. Risk of false negativity in samples with high-threshold cycle (i.e., low-viral load) value was a concern. Some researchers suggest adding few additional cycles to lower the chances of missing positive cases with low-Ct value. Lower limit of detection (LoD) of RT-PCR kits, that is, sensitivity of kits was another factor to consider. Thus, in a country like India, given the economic benefit and scarcity of resources, pooling strategy can be very effective, especially in low-prevalence areas and in low-risk contacts.

Introduction

The year 2020 took the whole world by storm on account of the emergence of a novel virus causing severe respiratory disease, which is believed to be due to a zoonotic spillover from wild animals (bats?). This condition first originated in the Wuhan town of China. The virus was found to belong to the family Coronaviridae which currently comprises 39 species of Corona viruses. The Coronaviridae Study Group (CSG) of the International Committee on Taxonomy of Viruses opined that the new virus is a sister clade to the prototype human and bat severe acute respiratory syndrome coronaviruses (SARS-CoV) and accordingly named it as SARS-CoV-2.1 On February 11, 2020, the World Health Organization (WHO) announced the name of the disease caused by the new virus to be"COVID-19," that is, Coronavirus Disease-2019.2 In a very
short time, COVID-19 became a major outbreak globally and on March 11, 2020, WHO declared it a pandemic, expressing grave concern about the alarming levels of spread and severity, and levels of inaction.\textsuperscript{2,3} Besides, the COVID-19 pandemic has also negatively impacted world economy.

Although in India disease progression was initially slow, the positive cases started exponentially rising by the end of March and beginning of April. The National Institute of Virology (NIV), Pune, with a biosafety level (BSL) 4 facility, was initially doing the laboratory diagnosis in India. Later on, other laboratories of Indian Council of Medical Research (ICMR) and functional Viral Research & Diagnostic laboratories (VRDLs) were roped in for diagnosis of COVID-19. Currently (June 21, 2020), around 992 laboratories (726 government and 266 private laboratories) are engaged in COVID-19 testing. A total of 69,50,493 tests have been conducted to date.\textsuperscript{4}

The basic reproduction number ($R_0$) of SARS-CoV-2 typically ranges from $2$ to $4$, implying that the number of secondary cases directly generated from an infectious index case in a fully susceptible population ranges from $2$ to $4$.\textsuperscript{5,6} However, it can vary across space and time, conditioned by factors such as adherence to measures like social distancing, masks, etc. Since the infected people shed virus even before the onset of symptoms, contact tracing, early detection of infected people, including asymptomatic cases, quarantine, etc. are key epidemiologic strategies for control of community spread.\textsuperscript{5} However, limited availability of testing kits, technical constraints and the high cost of tests, etc. are hindrances to achieve this target. Moreover, India being a densely populated country with more than 130 billion people, the challenge is much stiffer. Therefore, a pooled-sample testing protocol to screen large populations more rapidly and with limited resources may be an important strategy to increase screening capacity and accelerate testing for COVID-19.\textsuperscript{4}

The practice of “pooling of samples” dates back to 1943 when Robert Dorfman suggested the introduction of pooling of serum samples for syphilis screening in WW2 recruits. Since then, this practice has undergone multiple modifications. The pooling of RT-PCR samples has been found to be effective in screening HIV, chlamydia, malaria, influenza, etc.\textsuperscript{7} This practice involves the creation of pools by mixing of samples and performing the test on the “pool” as a single sample.\textsuperscript{8} If the result is negative, each constituent sample (of the pool) is regarded to be negative and no further individual testing is required. If the pool comes positive, each and every sample (of the pool) have to be further processed, according to the testing scheme (commonly tested individually). This leads to gross reduction in the number of reactions, depending on the number of samples pooled, that is, “size of the pool.”\textsuperscript{8}

### Objectives

a) To assess merits and demerits of sample pooling strategy for SARS CoV-2 diagnosis by RT PCR.

b) To analyze the factors influencing an effective sample pooling strategy.

### Materials and Methods

#### Literature Search Strategy

A search was performed to find research articles related to the pooling of samples for COVID-19 testing, which were published/accepted for publication from January to May 2020. Active search for articles was performed in the following databases: Medline, Embase, Google Scholar, CNKI, Wang Fang (the final two are primary databases for research in China—for abstract), etc. Search keywords were as follows: English MeSH keywords and Emtree terms. For example, [SARS-CoV-2 AND Pooling] OR [2019-nCoV ANDRT-PCR Pooling] OR [COVID-19 AND sample pooling] OR [new coronavirus AND RT PCR AND pooling] OR [Wuhan Coronavirus AND RT PCR AND Pooling] OR [Coronavirus AND RT PCR AND Pooling].

Gray literature search was also performed using the same keywords on Google Scholar, bioRxiv, and medRxiv (preprint servers) to capture the most recently published articles. WHO/CDC databases of publications on SARS-CoV-2 were also searched for relevant articles. Furthermore, related articles were also retrieved from the reference list and abstracts of published articles.

#### Inclusion and Exclusion Criteria

All the retrieved articles were screened for relatedness to the topics under study by both the authors independently. A consensus was drawn between both the researchers regarding inclusion and exclusion criteria. We considered studies that evaluated the outcome of sample pooling strategy in COVID-19 diagnosis compared with individual testing. Theoretical articles, commentaries, reports, and news articles were excluded. Nondiagnostic articles like economic and mathematical aspects of pooling of samples were also excluded. Studies without appropriate data essential in synthesis of results were excluded. Two authors selected the articles independently, according to the eligibility criteria, and final articles were unanimously selected by both authors.

#### Article Quality Evaluation:

The Newcastle-Ottawa quality assessment scale for cross-sectional studies was used. – Table 1 shows the results of this evaluation

#### Data Extraction

Data extracted from the selected articles include the following: author, setting/place of study, rate of positivity in that particular community/population, target gene for COVID-19 RT-PCR, false positivity and false negativity rate of the test kit (if available), optimal size of pool, pooling design, expected reduction in number of tests, etc. One of the authors performed the data extraction, while the other assessed the accuracy of the extracted data.

#### Results and Discussion

Sixty articles from different databases were primarily selected. Out of these, 19 were excluded as duplicates and of the remaining 41 articles, 13 were screened out by reading
the abstract, as they did not meet the inclusion criteria. Nine more articles were eliminated after reading the full text, due to nonavailability of necessary information. Of the remaining 19 articles, 10 were found to be pure mathematic modelling work, without any essential data for current review, and were removed. Finally, nine original research articles were selected for this review (► Fig. 1).

In all these articles, diagnosis of COVID-19 was done by real-time reverse transcriptase quantitative polymerase chain reaction (RT-qPCR). Detection of the SARS-CoV-2 specific targets like e-gene (for screening), s-gene, ORF1ab (RdRp), and ORF1bin combination implies a positive test.

Optimal Pool Size

Three studies8–11 observed that larger pool size, of around 25 to 30 samples, can accurately detect a positive sample. Yelin et al arranged the negative samples in the pools in ascending numbers: 1, 3, 7, 15, 31, and 63 and one positive sample in each pool. Here, positive samples were consistently detected up to a pool size of 32. However, false negativity was detected in 10%.8 Lohse et al showed 30 samples per pool can increase test capacity with existing resources and detect positive cases with sufficient diagnostic accuracy.8 The study by Deckert et al observed accurate results up to group size of 25 in low-to-moderate prevalence.11

Another three studies performed group testing by taking relatively small pool sizes of 5 to 10 samples.12–14 Abdalhamid et al observed that positive samples (detected by individual testing) remain positive even if pooled with 4 negative samples. They created 21 pools with each pool containing 5 samples (1 known positive + 4 known negative). They also observed that in low-prevalence setting, pool size can be increased (► Table 2).15 Hogan et al pooled 9 to 10 samples and screened 292 pools to detect community transmission of SARS-CoV-2. They reported only one false positive.13 Gupta et al took extracted RNA eluates and randomly pooled in groups of 2, 4, 6, 8, and 16. Best result, when compared with individual testing, was observed in 2, 4, and 8 sample pools. Thirty four pools out of 35, showed concordance with individual testing, that is, 95.4% pools correctly identified the positive samples.14

Two simulation studies parameterized various pool sizes and multistage testing schemes and compared their efficiency at different prevalence rates.15,16 Eberhardt et al found that three-stage schemes performed optimally for prevalence rates up to 12%, and that initial pool sizes of 16 samples were best for prevalence rates up to 3.5% and pools of 9 samples for rates between 3.5% and 12% (P9S3, improvement factor 3.8 to 1.5). For prevalence rates between 12 and 30%, two-stage testing with pools of three samples performed best (► Table 2).15 Shani-Narkiss et al compared one-time pooling and multiple pooling steps and built the protocol for optimal batch size for pooling samples, for any given p (expected frequency of positive samples out of all samples) (► Table 2).16 Sinnott-Armstrong et al, instead of mentioning a specific pool size/group size, adopted strategy of pooling 8 rows and 12 columns of a 96-well plate and reduced the number of reaction from 96 (in individual testing) to 20 reactions. They took 36 reaction to find 5 positive samples (in 96-well setting).17

Effect of Prevalence/Rate of Positivity on Pool Size

Prevalence of the disease and the rate of positivity, reflects the infection level of population in that particular area. This factor considerably affects the optimal pool size. Six of the included studies have shown that lower the prevalence of the disease, greater will be the number of samples pooled together.10–12,15,17 Lohse et al found sufficient diagnostic accuracy up to pool size of 30 with prevalence 1·93% and rate of positivity 4·24% during that period.10 ► Table 2 clearly shows that ‘pool size’ can be increased proportionately as positivity rate/prevalence decreases.12,15,16 An inverse relation of group size to rising prevalence from low to high was observed in a simulation study by Deckert et al.11 Sinnott-Armstrong et al observed that at prevalence < 2%, row and column pooling on 384-well plates performs best; between 2 to 10%, row and column pooling on 96-well plates performs best; and at > 10% prevalence, four-way pooling of wells performs best.17

Effect of Pooling on the CT Value of the Positive Result

Four studies noticed increase in the threshold cycles (Ct) value upon pooling compared with the Ct detected during individual testing. Yelin et al observed that as the number of negative pooled samples increases (from 1:2 to 1:32), the amplified RNA reaches the threshold later, as expected from a diluted sample; but majority pools (up to 32-sample pools) reached the threshold, only one of the ten tested replicates, did not cross the threshold in pools of 32.8 Similar results were obtained by Abdalhamid et al on comparing the

<table>
<thead>
<tr>
<th>Studies (First author)</th>
<th>Selection</th>
<th>Comparability</th>
<th>Outcome</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yelin et al8</td>
<td>5****</td>
<td>**</td>
<td>***</td>
<td>9*</td>
</tr>
<tr>
<td>Lohse et al10</td>
<td>5****</td>
<td>**</td>
<td>**</td>
<td>9*</td>
</tr>
<tr>
<td>Abdalhamid et al12</td>
<td>5****</td>
<td>**</td>
<td>***</td>
<td>9*</td>
</tr>
<tr>
<td>Hogan et al13</td>
<td>5****</td>
<td>**</td>
<td>**</td>
<td>9*</td>
</tr>
<tr>
<td>Gupta et al14</td>
<td>****</td>
<td>**</td>
<td>**</td>
<td>7*</td>
</tr>
<tr>
<td>Eberhardt et al15</td>
<td>5****</td>
<td>**</td>
<td>***</td>
<td>9*</td>
</tr>
<tr>
<td>Shani-Narkiss et al16</td>
<td>****</td>
<td>**</td>
<td>**</td>
<td>6*</td>
</tr>
<tr>
<td>Deckert et al11</td>
<td>****</td>
<td>**</td>
<td>**</td>
<td>8*</td>
</tr>
<tr>
<td>Sinnott-Armstrong et al17</td>
<td>****</td>
<td>**</td>
<td>**</td>
<td>6*</td>
</tr>
</tbody>
</table>
Ct between the original and pooled COVID-19 positive samples and observed that there was increase in Ct value in 20 out of 21 pools from the original Ct value. But all 21-pooled specimens were positive within a range of −1.1 Ct to 5.09 Ct difference from the original samples. Gupta et al observed overall mean Ct value of individual testing as 32.68, while for pooled testing it was 34.24, that is, an increase in Ct by 1.56 but very much within the reporting criteria of being called as positive for the kit used (i.e., < 40). However, Lohse et al observed lower Ct values in some retested positive individual samples. They hypothesized that “the lower Ct values of pools than that of single samples were because of the carrier effect of the higher RNA content.” But overall, for 4 to 30 samples per pool, Ct values of positive pools were between 22 and 29 (E-gene) and between 21 and 29 (S-gene), which were higher compared with Ct values of individual testing of positive sample. Thus, this observation does not rule out the chance of missing the borderline
positive samples with Ct value on the higher side in pooled testing. This is the major disadvantage of group testing. However, adding a few additional PCR cycles could be considered as a means to increase detection rate of low-viral load samples.

**Reduction in Number of Reactions**

All the included studies observed significant reduction in the number of total tests, depending on pool size and prevalence. Abdalhamid et al observed expected number of tests per individual to be < 0.44 with optimal pool size of 5 and prevalence rate at 5% (►Table 2). Sinnott-Armstrong et al also found maximum reduction in tests with 384-well plates, provided the prevalence is low (< 2%). Eberhardt et al defined the term “improvement factor” to compare the performance of different group testing schemes with individual testing and observed its inverse relation with prevalence (►Table 2). Shani-Narkiss et al also showed how fraction of test (required) decreases with increase in pool size and decrease in p (frequency of positivity) (►Table 2). Deckert et al also observed 56% to 93% less tests in low-to-moderate prevalence settings and with group sizes up to 25. In two included studies, extracted RNA from the individual samples are pooled together; however, pooling can also be done from patient samples before RNA extraction, leading to further reduction of consumables, manpower and time.

**Effect of Sensitivity and Specificity of the Assay**

Abdalhamid et al opined that effective pooling strategy can be achieved only with test assays with high-sensitivity and specificity. The molecular assay used in their study had a sensitivity and specificity of 100% with lower limit of detection of RNA copies of 1 to 3 copies/ul. A study by Hanel et al showed that false-negative factor for the pooled strategy increases with the false-negative rates of the test assay, and this can be lowered by using replicates. As sensitivity and specificity of different RT-qPCR kits and protocols vary with make; it is suggested that pooling may require validation for each specific setting.

**Effect of Dilution**

Another drawback of pooled testing is that it may reduce the test’s sensitivity due to pooling dilution. If one of the samples in the pool is infected, the viral load will be diluted due to pooling with negative samples. However, a study by Theagarajan has shown that pool size of 32 still provides a very high sensitivity after pooling dilution, even when a single sample is infected.

Included studies are summarized in ►Table 3.

**Conclusion**

It is clear from above that pooling strategy can be effectively applied for the detection of SARS-CoV-2 by RT-qPCR, as supported by other studies. Optimal pool size has to be determined by the level/rate of infection in that particular population. Currently, positivity rate (p) in India is 6.1; although due to diversity in population distribution, there is great variance in this rate from one state to another. While in some cities like Mumbai, Delhi, Ahmadabad, etc. positivity...
rate is very high, some states still have very low p, where pooling strategy can play an important role. Greater the optimal pool size, greater will be the reductions in reactions, thereby saving the scarce diagnostic resources in this time of crisis. However, diagnostic assays used for pooling should have significantly low-false positivity and false negativity. Nonavailability of analytical parameters like LoD, sensitivity, specificity, etc. in most SARS-CoV-2 PCR kits is a serious gap in current studies. These have implications in pooling strategies, especially in low-viral load samples. Hence, it is an important scope for future analysis/study on this topic.

**Acknowledgment**

We acknowledge AIIMS Rishikesh for allowing us to carry out the study. We are also thankful to the authority of AIIMS Rishikesh for providing us with the resources to conduct this study.

**References**


**Table 3** Overview of the nine studies included in the scoping review

<table>
<thead>
<tr>
<th>Study, setting, and rate of positivity</th>
<th>Study design, time horizon, target gene</th>
<th>Kit parameters</th>
<th>Pool size</th>
<th>Outcome and comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yelin et al® Israel Institute of Technology, Haifa p: not mentioned</td>
<td>67 negative and 5 positive samples pooled to 6 different pool sizes. Target: e-gene</td>
<td>NM</td>
<td>2, 4, 8, 16, 32 and 64 Opti: 32</td>
<td>32 samples pooling can be useful with 10% false-negative result risk.</td>
</tr>
<tr>
<td>Lohse et al® Hamburg Germany p: 4.24% in institute and 1.93% overall</td>
<td>30, 10 and individual samples sequentially. Target: e- and s-gene</td>
<td>NM</td>
<td>30, 10 Opti: 30</td>
<td>Test capacity increased without loss of diagnostic accuracy.</td>
</tr>
<tr>
<td>Abdalhamid et al® Nebraska Public Health Laboratory, USA p: 5% (prevalence rate)</td>
<td>Initial 5 samples pool followed by community validation (60 samples in 12 pools)</td>
<td>LOD: 1–3 copies/ul; Sen: 100%</td>
<td>3 to 10; experimented size 5 Opti: 5</td>
<td>Reservation of resources with 69% increase in testing capacity</td>
</tr>
<tr>
<td>Hogan et al® Stanford, USA p: less than 1%</td>
<td>Retrospective study on NPS and BAL. Screening for e-gene (in pool) and RdRp (individual)</td>
<td>NM</td>
<td>9 and 10 Opti: NM</td>
<td>Increase testing throughput, may miss cases in low-risk setting.</td>
</tr>
<tr>
<td>Gupta et al® New Delhi, India p: not mentioned</td>
<td>Extracted RNA (240) into 35 pools (8 each) and tested in groups and individually. Gene: e and RdRp</td>
<td>NM</td>
<td>Pool of 8 samples Opti: 8</td>
<td>8 sample pool is reliable &amp; detect up to single positive sample</td>
</tr>
<tr>
<td>Eberhardt et al® Bonn, Germany p: various prevalence rate</td>
<td>Multistage group testing schemes with improvement factor depending on different prevalence rates</td>
<td>NM</td>
<td>3 stage with various probabilities</td>
<td>With increasing p, optimal pool size decreases (Table 2)</td>
</tr>
<tr>
<td>Sinnott-Armstrong et al® Stanford, USA Low prevalence</td>
<td>i) Single dimension 4 pools ii) 96-well plate: column X row iii) 384-well plate column X row</td>
<td>NM</td>
<td>4 well one-dimension, 96-well (8 × 12) plate and 384-well (16 × 24) plate. Opti: 96-well plate</td>
<td>Useful in low-prevalence countries, especially in absence of a calculated “p”</td>
</tr>
<tr>
<td>Deckert et al® Heidelberg Univer, Germany. p: various prevalence</td>
<td>Comparison of simulation of two groups</td>
<td>NM</td>
<td>Multiple pool sizes</td>
<td>Modeling study clearly indicating advantage of pooling approach,</td>
</tr>
<tr>
<td>Shani-Narkiss et al® Hebrew Univ. Jerusalem Various theoretical prevalence considered</td>
<td>Modeling study: repeated pooling in several stages and one-stage pooling</td>
<td>NM</td>
<td>Various pool sizes (mathematical consideration). Opti: 64</td>
<td>Multistaging pooling is useful for big advance laboratories; one stage for other laboratories</td>
</tr>
</tbody>
</table>

Abbreviations: BAL, bronchoalveolar lavage fluid; NM, not mentioned; NPS, nasopharyngeal swab; p, rate of positivity; Opti, optimal pool size; Sen, sensitivity.
19 Nguyen NT, Aprahamian H, Bish EK, Bish DR. A methodology for deriving the sensitivity of pooled testing, based on viral load progression and pooling dilution. J Transl Med 2019;17(1):252
20 Gollier C, Gossner O. Group testing against Covid-19. Available at: https://hal.archives-ouvertes.fr/hal-02550740. Accessed July 31, 2020
21 Tauer M. Rapid, large-scale, and effective detection of COVID-19 via non-adaptive testing. Available at: https://www.biorxiv.org/content/10.1101/2020.04.06.028431v1. Accessed on June 1, 2020