

Optimal sample pooling: an efficient tool against SARS-CoV-2

Saurabh Goyal^a, Priti Bist^a, and Rakesh Sharma^b

^a*Digital experimentation team, American Express, Gurugram, Haryana, India*

^b*Comofi Medtech Pvt Ltd, Bangalore, Karnataka, India*

Abstract

The SARS-CoV-2 pandemic situation has presented multiple imminent challenges to the nations around the globe. While health agencies around the world are exploring various options to contain the spread of this fatal viral infection, multiple strategies and guidelines are being issued to boost the fight against the disease. Identifying and isolating infected individuals at an early phase of the disease has been a very successful approach to stop the chain of transmission. But this approach faces a practical challenge of limited resources. Sample pooling solves this enigma by significantly improving the testing capacity and result turn around time while using no extra resources. However, the general sample pooling method also has the scope of significant improvements. This article describes a process to further optimize the resources with optimal sample pooling. This is a user-friendly technique, scalable on a national or international scale. A mathematical model has been built and validated for its performance using clinical data.

Introduction

The rapid rise in SARS-CoV-2 positive cases poses a significant health threat to the world. In India, the case count is surging as Government has announced significant relaxations in curbs for Unlock 1.¹

Identifying and isolating the positive cases in the very early phase has been a major successful approach to stop the spread of the disease.^{2,3} To identify such cases, the World Health Organization has stressed on multiple occasions the significant role of sample testing.^{4,5} Yet unfortunately, in the testing strategies of majority of the nations, either only symptomatic cases are included or are heavily favoured.⁶ With such approach, the patient count tracker will always find itself trailing the steps of the viral spread. Multiple scientific studies state that the disease can lay asymptomatic for the first 2 to 3 weeks, immediately followed by the period of heavy viral shedding just as symptoms onset.⁷⁻¹³ Thus, testing all of the suspected individuals will provide an ideal solution. However, a major challenge in implementing such approach is resource limitation.

Many strategies have been identified to release the constraint of resource limitation. One of such tried and tested approach is sample pooling.¹⁴⁻¹⁹ Indian Council of Medical Research (ICMR) and other world health agencies have also suggested sample pooling as a major strategy to increase the scope of the test.^{20,21} This strategy helps in multiple ways, it reduces both cost and result time.

Sample pooling can be achieved primarily in two ways: Repeated pooling and one-time pooling.¹⁹ In repeated sample pooling, a sample pool which has tested positive is further broken down into sub-pools and the testing process is repeated till each of the individual samples with positive SARS-CoV-2 infection is identified. This method turns out to be more time and cost efficient as it can be designed by using algorithms like binary search tree.¹⁹ However, this method has serious practical limitations as – a) it is a complex procedure and requires training to implement and b) it shows significantly higher efficiency when the initial pool size is high. In contrast, during one-time pooling, only one level of pooling is done. If a pool tests positive, in the next step, each individual sample of the pool is tested for the infection. Hence, one-time pooling is a more practical approach than the repeated pooling.

While the current guidelines from ICMR state that up to 5 samples can be pooled,²⁰ multiple studies have confirmed that the pooling size of up to 8 does not harm the specificity and the sensitivity of the test.¹⁶ The latest guideline has left it up to the lab to decide the optimal pool size. Multiple tools have been reported in the scientific community to find the optimal pool size.¹⁵ While most of such tools consider multiple parameters to arrive at the optimal pool size, complicated procedures avert the practical use.

The current paper proposes a practical approach to find the optimal pool size. This approach is simple, easy to implement on a national scale and is adaptable to the stage of the viral spread.

Method

Testing facilities are either already designated to test samples from a specific geography or can be easily mapped. This correlation ensures that the prevalence rate of the lab remains a continuous function. Sample data verifies this hypothesis. The temporal variance in the prevalence rate is significantly lower for a testing facility level than at a country or even at a state level where the variance is extremely high.²³ Hence, in order to determine the current prevalence rate and to forecast the same, the metrics should be calculated for individual testing facilities. The determination of sample pool size for each lab using its prevalence rate would yield desired efficiency and be easy to implement.

Following mathematical model describes the gain factor of the sample pooling against individual testing of each sample. Involved parameters are,

Current prevalence rate, the probability of a sample being positive is forecasted using the actual prevalence rates of the last one week, p

Samples to be tested, n

Sample pool size, x (min – 2, max – 8)

No of tests without sample pooling, $N_{\text{normal}} = n$,

No. of tests after sample pooling,

$$N_{\text{pool}} = \frac{n}{x} + n * \{1 - (1 - p)^x\} \quad (1)$$

Gain factor,

$$G = N_{\text{pool}} / N_{\text{normal}} = 1 / \left[\frac{1}{x} + \{1 - (1 - p)^x\} \right] \quad (2)$$

Using equation (2), gain factor values for different sample pool size and varying prevalence rate are simulated (Table 1).

GAIN FACTOR		Pool Size						
		2	3	4	5	6	7	8
Prevalence Rate	1%	1.9	2.8	3.5	4.0	4.4	4.7	4.9
	3%	1.8	2.4	2.7	2.9	3.0	3.0	2.9
	5%	1.7	2.1	2.3	2.3	2.3	2.2	2.2
	10%	1.4	1.7	1.7	1.6	1.6	1.5	1.4
	20%	1.2	1.2	1.2	1.1	1.1	1.1	1.0

Table 1. Gain factor simulation results for different sample pool size and varying prevalence rates

Min	Max	Result	Pool size
0%	1%	4.0	5
1%	3%	2.9	5
3%	5%	2.3	5
5%	10%	1.7	4
10%	20%	1.2	4
20%	100%	Do not Pool	Do not Pool

Table 2. Decision matrix for optimal pool size selection

In the situation of prevalence rate of 1% or lower, though higher sample pool size yields better gain factors, yet the risk of sample dilution outweighs the miniscule gains. Thus, the sample pool size of 5 is suitable during lower prevalence rates, less than 5%, while for higher prevalence rates, more than 5%, pool size of 4 provides the best efficiency. However, for prevalence rates above 20%, sample pooling should not be used.

Discussion and future work

Strategizing to have a common sample pool size across the nation would not yield optimum results as the variance of prevalence rates is extremely high. Hence, sample pool size should be decided individually for a testing facility using the prevalence rates recorded by the same lab. Our results show that a period of 7 days can be used to forecast the prevalence rate of the next day. This prevalence rate can then be looked up on the decision matrix table to arrive at the optimal sample pool size. However, sample from the following categories should be excluded from the process:

1. A pool or an individual sample which has already been tested positive
2. Special testing scenarios, such as, retesting or testing of a specific group which is bound to be having much higher risk than the normal conditions.

Though the mathematical model has been tested with the clinical data, a clinical study to test the performance is being planned.

Supplementary Information

Prevalence rate variation of a lab in clinical setting.

References:

- [1] Ministry of Home Affairs (MHA), The Government of India, Order No. 40-3/2020-DM-I (A) dated 30th May, 2020, Available from <https://www.mha.gov.in>
- [2] Cowling, Benjamin J., et al. "Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study." *The Lancet Public Health* (2020).
- [3] Bi, Qifang, et al. "Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study." *The Lancet Infectious Diseases* (2020).
- [4] World Health Organization. Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance, 2 March 2020. No. WHO/COVID-19/laboratory/2020.4. World Health Organization, 2020.
- [5] World Health Organization. Laboratory testing strategy recommendations for COVID-19: interim guidance, 22 March 2020. No. WHO/COVID-19/lab_testing/2020.1. World Health Organization, 2020.
- [6] Ministry of Health and Family Welfare, Government of India. Strategy for COVID19 testing in India (Version 5, dated 18th May 2020). Available from: <https://www.mohfw.gov.in/>
- [7] He, Xi, et al. "Temporal dynamics in viral shedding and transmissibility of COVID-19." *Nature medicine* 26.5 (2020): 672-675.
- [8] Zou, Lirong, et al. "SARS-CoV-2 viral load in upper respiratory specimens of infected patients." *New England Journal of Medicine* 382.12 (2020): 1177-1179.
- [9] Santarpia, Joshua L., et al. "Transmission potential of SARS-CoV-2 in viral shedding observed at the University of Nebraska Medical Center." *MedRxiv* (2020).
- [10] To, Kelvin Kai-Wang, et al. "Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study." *The Lancet Infectious Diseases* (2020).
- [11] Pan, Xingfei, et al. "Asymptomatic cases in a family cluster with SARS-CoV-2 infection." *The Lancet Infectious Diseases* 20.4 (2020): 410-411.
- [12] Kimball, Anne, et al. "Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility—King County, Washington, March 2020." *Morbidity and Mortality Weekly Report* 69.13 (2020): 377.
- [13] Al-Tawfiq, Jaffar A. "Asymptomatic coronavirus infection: MERS-CoV and SARS-CoV-2 (COVID-19)." *Travel medicine and infectious disease* (2020).
- [14] Hogan, Catherine A., Malaya K. Sahoo, and Benjamin A. Pinsky. "Sample pooling as a strategy to detect community transmission of SARS-CoV-2." *Jama* (2020).
- [15] Abdalhamid, Baha, et al. "Assessment of specimen pooling to conserve SARS CoV-2 testing resources." *American Journal of Clinical Pathology* 153.6 (2020): 715-718.
- [16] Gupta, Ekta, et al. "Pooled RNA sample reverse transcriptase real time PCR assay for SARS CoV-2 infection: a reliable, faster and economical method." *medRxiv* (2020).
- [17] Bertolotti, Paolo, et al. "Pooled Sample Testing for SARS-CoV-2." (2020).
- [18] Yelin, Idan, et al. "Evaluation of COVID-19 RT-qPCR test in multi-sample pools." *MedRxiv* (2020).
- [19] Shani-Narkiss, Haran, et al. "Efficient and Practical Sample Pooling High-Throughput PCR Diagnosis of COVID-19." *medRxiv* (2020).
- [20] Ministry of Health and Family Welfare, Government of India. Advisory on feasibility of using pooled samples for molecular testing of COVID-19 (dated 14th April, 2020). Available from: <https://www.mohfw.gov.in/>
- [21] Ministry of Health and Family Welfare, Government of India. Guideline for RT-PCR based pooled sampling for migrants/returnees from abroad/green zones (dated 14th May, 2020). Available from: <https://www.mohfw.gov.in/>
- [22] PBDS University of Health Sciences, Rohtak, Haryana, India. COVID-19 testing Bulletin, Available at <http://www.uhsr.ac.in/>
- [23] No uniformity in positivity rate of Covid samples at laboratories, Tribune New Services India, May 8. Available at <https://www.tribuneindia.com/news/haryana/no-uniformity-in-positivity-rate-of-covid-samples-at-laboratories-82320>