

**SCIENTIFIC WHITEPAPER:**

**GUIDANCE ON  
POOLING OF  
SAMPLES FOR  
SARS-COV-2**



Daniel Westreich <sup>1,2</sup>

Karen Diepstra <sup>2</sup>

Ryan Max <sup>2</sup>

<sup>1</sup>Gillings Center for Coronavirus Testing, Screening, and Surveillance and

<sup>2</sup>Department of Epidemiology, UNC-Chapel Hill

Final as of 14 December 2020. As information continues to emerge about SARS-CoV-2 and COVID-19, all contents are subject to update and change. Contact [djw@unc.edu](mailto:djw@unc.edu) with questions.

Thanks to: Jessie K. Edwards, Michael Hudgens, Penny Gordon-Larsen, Christopher D. Pilcher, Jonathan J. Juliano, Jessica Lin.

## EXECUTIVE SUMMARY

Pooled testing, also known as group or batch testing, is a process in which portions of individual samples are combined into a single pool, which is then tested for a biomarker. If the pool is negative, each of the samples within that pool is considered negative and no further testing is required. If the pool is positive, the individual samples that contributed to it are tested to determine which of them are positive. Compared to individually testing each sample, pooled testing strategies can increase the efficiency, speed, and positive predictive value of diagnostics for case identification. These strategies (in some form) have been in practice since 1943 and have become more powerful with conceptual and technological advances in the interim. However, simple pooling approaches are

straightforward, can lead to large gains in efficiency, and can be implemented in nearly any laboratory setting with minimal preparation.

Pooled testing is generally most useful when the prevalence of positive samples among those being tested is low; the usefulness of the strategy decreases as the proportion of samples that are positive increases. Pooled testing also relies on an underlying assumption that after samples are combined and tested for the presence of a biomarker of infection, a negative result indicates that all constituent samples are negative. However, whether this assumption is reasonable depends on the proportion of samples that would have tested positive under individual testing but no longer do so under pooled testing. Pooled testing decreases the overall sensitivity of the testing process, as the concentration of the biomarker in a positive sample may no longer be above the assay's limit of detection after being pooled with negative samples. At the same time, pooling typically increases the positive predictive value of testing compared to individual testing because pooling strategies result in repeat testing of positive specimens, and thus an increase in diagnostic specificity. The benefits of pooling should be considered alongside the loss of sensitivity from dilution, as well as the increased complexity in laboratory procedures and potential for contamination that is introduced by pooling.

Presented here are the results of a free, web-based R Shiny app developed to determine the optimal pooling size for SARS-CoV-2 RT-PCR testing based on assay sensitivity and specificity, the underlying prevalence of SARS-CoV-2 among samples tested, and making the assumption that we will not allow greater than 20% dilution-related loss of sensitivity compared to individual testing. We examined two pooling strategies, one simpler and one more complex. For both strategies and at all prevalence levels examined (between 0.1% to 20%), efficiency and positive predictive value improve under pooling relative to individual testing. Further, while total sensitivity is reduced

by pooling, every pooling scenario we examined is able to identify more true positive cases per assay compared to individual testing.

Pooling can serve as a straightforward way to increase the efficiency of testing for SARS-CoV-2, with relatively modest downsides. It should be considered as a potential testing strategy moving forward.

# I. BACKGROUND ON POOLED TESTING

## HISTORY AND PARADIGMS

**Pooling for Case Identification.** Pooled testing (alternatively known as group testing or batch testing) of populations for the identification of infected individuals was popularized (or perhaps invented) by Robert Dorfman, who in 1943 demonstrated the potential to save time and resources on syphilis testing for military recruits [1]. Dorfman showed that labs could take a fraction of each blood sample, combine those samples into pools of a fixed number of individuals, and test the pools. If the pools were negative, the individuals in those pools could be assumed negative, while if a pool was positive further testing could be done to determine which individual(s) were positive. Dorfman demonstrated substantial reduction in the total number of tests that had to be run using pooled instead of individual testing.

**Pooling for Prevalence Estimation.** Pooling of samples can also be used to estimate the prevalence proportion of infection in a population without testing each individual separately. In practice, this resembles the first step of Dorfman's staged testing algorithm, but without the follow-up testing of individuals from positive pools. This approach is suited to most applications where the identification of individual

cases is not a public health or research priority, and is particularly useful in cases where the assay being used has a quantitative output that can be interpreted as the level of infection in a pool (such as antibody concentration or viral titer), or where prevalence is too low for individual testing to be practical for prevalence estimation[2].

## II. WHY CONSIDER POOLING FOR SARS-COV-2?

Pooling has numerous advantages compared to individual testing. We explain these advantages and disadvantages briefly.

### ADVANTAGES OF POOLING

**Efficiency**, the number of results obtained per test run, can be increased, especially when test positivity is low. Note that this may allow us to test more widely in asymptomatic populations with a lower prevalence.

**Positive predictive value**, the probability that a positive test result is a true positive, can be increased. This is because a truly negative sample will not be declared false-positive until it has been through multiple rounds of testing in pooling: if errors are independent between testing rounds, the probability of a false-positive result will become much smaller and thus positive predictive value will increase (see below for more detail).

**Throughput**, the number of samples that can be processed over a given time period, can be increased. If a lab can process only 100 samples a day, and has one thousand samples to process, individual testing would require 10 days – whereas 10:1 pooling (approximately optimal at a prevalence of 0.01) can reduce the total time to at most 3

days: 1 day to test 100 pools of 10; 1 day to test the individual samples within any positive pools (around  $10 \times 10 = 100$  samples); and 1 day for additional lab-related logistics or testing of any additional individual samples.

## DISADVANTAGES OF POOLING

**Loss of sensitivity** is a disadvantage of pooling. Optimal pooling sizes for efficiency mean that (typically) at most one sample in a pool is positive; thus, in a pool of size  $N$  with a single positive sample, the signal of that sample (e.g., the viral load measured as a concentration) will be diluted by a factor of  $N$ . If this brings the viral load in that sample below the limit of detection for the assay, this positive sample will be missed.

**Increased lab complexity.** Pooling requires additional effort from laboratory personnel (or the use of a pooling robot); in addition to taking more time to process each individual sample (though reduced time overall, see *Throughput* above), this may risk cross-contamination of samples. Contaminated samples will compromise the *positive predictive value* advantage mentioned above.

## III. UNDERSTANDING POOLING: CONCEPTUAL OVERVIEW

In this section we will discuss several of the pooled testing strategies that have been used historically, some of which are currently being used to detect SARS-CoV-2.

### POOLED TESTING STRATEGIES

*(Figures in this section from Westreich et al. J Clin Microbiol. 2008)*

**Two-stage hierarchical (D2).** The testing strategy described by Dorfman in his 1943 paper is a two-stage hierarchical testing scheme, sometimes called “minipooling.” First, parts of the individual test samples are combined to form a pool; the pool is tested; and then individuals from positive pools are tested. For example, in a group of 10 individuals, a pool of 10 would be made. In a group of 100 individuals, a single pool of 100 samples, or 10 pools of 10 samples, could be made and tested. This strategy is visually represented in Figure 1.

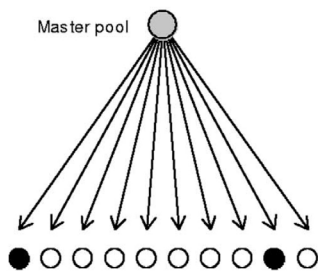


Figure 1: two-stage pooling (D2).

**Three-stage hierarchical (D3).** Three-stage hierarchical testing is an extension of the motivating example and scheme provided by Dorfman, with an additional step. After individual samples are combined into minipools, these minipools are combined into higher level pools. The largest pools are tested; if positive, then the constituent minipools of any positive master pools are tested; and finally the individual samples of any positive minipools are tested[3]. For example, 100 individuals could be tested by creating 10 minipools comprising 10 individuals each, and a master pool comprising all 100 samples from those 10 minipools. This strategy is visually represented in Figure 2.

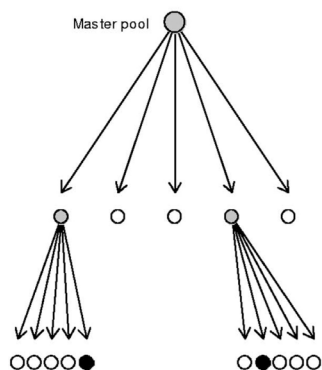


Figure 2: three-stage pooling (D3).

**Higher-order hierarchical testing (DN).** Hierarchical testing can be extended to more than three stages as well.

**Array Pooling.** Array pooling (typically, *square* array pooling, sometimes abbreviated A2m) requires that samples be arranged (literally or conceptually) on a grid, like a chessboard. Pools are created for each row and each column, and individuals whose row and column pools both test positive are then tested individually [4]. A master pool may also be created of all samples, depending on the acceptable level of dilution. A conceptual diagram is shown in Figure 3.

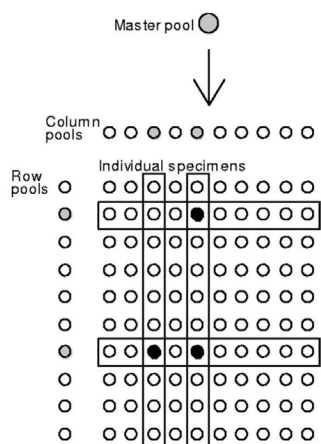


Figure 3: square array pooling (A2m).

**Other designs.** Beyond n-stage and array pooling, other pooled testing designs are being considered for the detection of SARS-CoV-2 infection. Shental et al. have proposed a combinatorial testing design which



removes the need for multiple rounds of testing. Each patient sample is split among multiple pools, and positive individuals are identified based on the combination of pools with positive results and the knowledge of which samples were present or absent in each pool[5]. Mutesa et al. discusses a strategy being tested in Rwanda similar in principle to array pooling, but increasing the number of dimensions[6]. These offer a potential advantage over the hierarchical strategies, but are best suited to sample populations with low prevalence (~1% or even lower).

We note that while all pooling strategies increase complexity in the lab compared to individual testing, array pooling as well as strategies listed under “Other designs” may increase complexity beyond what is practicable in lab settings without use of specialized automated pooling systems (“pooling robots”, e.g. epMotion 5070 robot (Eppendorf, Hamburg, Germany)).

## IV. UNDERSTANDING POOLING: TERMS AND METHODS

### DEFINITION/EXPLANATION OF TERMS

**Optimal pooling strategy.** The optimal pooling strategy, for a given set of parameters, is here chosen as the one that has the best value for efficiency (see below). We acknowledge that some investigators might wish to optimize the pooling process for something other than efficiency: for example, it might be preferable to accept slightly lower efficiency if it meant obtaining a higher positive predictive value. Here we do not address this possibility, but it is a possible subject for future work.

**Efficiency.** While we briefly defined efficiency above, here we elaborate. The primary benefit of pooled testing is the substantial increase in *efficiency in use of test kits* that can be gained compared to individual testing, especially if the population undergoing testing has a low prevalence of the disease of interest. There may be efficiency gains in turnaround time, as well, but we address this issue below (*Time to results*, and in Section VIII, first paragraph).

The efficiency of a testing strategy is defined as the average number of test kits used per result obtained. Individual testing always has an efficiency of 1 – the average number of tests used is 1 per person (or 1 per result). For example, individual testing of a population of 1,000 persons will use 1,000 tests (efficiency = 1). However, if pooling strategy A uses 500 tests to obtain results for the 1,000 persons and strategy B only 100 tests, efficiency will be 0.5 and 0.1, respectively. In this measurement of efficiency (tests per person), lower values are better.

Alternatively, we can describe efficiency in terms of the *number of people who can be expected to be screened per test used*: this is simply the reciprocal of the efficiency number above. If we can screen 1000 people with 500 tests, then on average we are screening 2 people for every test. In this measurement of efficiency (people screened per test), higher values are better.

Efficiency changes primarily with prevalence of the condition for which we are testing, but can also change with assay sensitivity and specificity, as well as anticipated dilution effects from the pooling itself. The R Shiny app (visit [gillingscovid19.unc.edu](https://gillingscovid19.unc.edu) and search “pooling” to find the research app) and an older web calculator ([bios.unc.edu/~mhudgens/SARS-CoV-2.pooling.home.html](https://bios.unc.edu/~mhudgens/SARS-CoV-2.pooling.home.html)) allow calculation of efficiency for arbitrary input parameters, and results from the two tools are identical.

**Case identification efficiency.** In addition to gains in efficiency in use of test kits, pooled testing can increase the expected number of true

positive cases detected per test (or per 1000 tests) used. This is true despite pooling reducing sensitivity [7]. In a simulation study, Cleary et al. examined the outcome “total recall” defined as the total number of positive individuals identified by a testing strategy [8]. Depending on the pooling strategy and disease prevalence in the simulated population, pooling in the context of limited test kit availability could identify as many as 20 times the number of true positive cases compared to individual testing [8], despite pooling-related losses in sensitivity.

Expected number of true positive cases detected per 1,000 tests is a function of the true prevalence in the population, and the efficiency and sensitivity of the testing algorithm (see below for more on sensitivity). The expected number of true positive cases detected per test used can be estimated as  $(1/\text{efficiency}) \times (\text{true prevalence of disease}) \times (\text{sensitivity of pooled testing})$ , where efficiency is measured as tests/person. In the individual testing scenario, efficiency is equal to 1, and diagnostic sensitivity (defined below) is the same as the sensitivity of the assay itself. We discuss sensitivity of pooled testing immediately below.

**Diagnostic test performance.** Here we presume little familiarity with clinical epidemiology methods, and so first explain sensitivity and specificity so that we can then explain how pooled testing leads to increases in positive predictive value and decreases in sensitivity due to dilution.

**Sensitivity and specificity.** Broadly, sensitivity is the ability of a diagnostic test or other tool to correctly identify true positive samples or individuals, and is strongly analogous to “power” in randomized trials settings. One minus the sensitivity is the false negative probability, sometimes called Type II error. Specificity is the ability of a diagnostic test or other tool to correctly identify true negative samples or individuals; the complement of specificity is the probability of a false positive, sometimes called a Type I error.

An important distinction, described by Saah and Hoover in 1997, exists between analytical and diagnostic sensitivity, and likewise between analytical and diagnostic specificity [9]. The authors defined *analytical sensitivity* as an “assay’s ability to detect a low concentration of a given substance in a biological sample,” which may be referred to as the lower limit of detection (LLD). The analytical sensitivity, or LLD, is typically determined by identifying the lowest concentration at which some proportion (for example, 95%) of known positive specimen are identified as positive by the assay [10,11].

In contrast, *diagnostic sensitivity* is a proportion defined as the percentage of persons with disease in a population undergoing testing that are *ultimately classified correctly as having that disease*. Diagnostic sensitivity is affected by the assay’s analytical sensitivity, in addition to other factors. For example, if a person who truly has COVID-19 receives inadequate nasal swabbing that does not capture any SARS-CoV-2 RNA, this individual’s nasal swab sample will likely test negative and the individual incorrectly classified as not having COVID-19 based on this result. This is not a failing of the assay, but rather of the cascade of events prior to the assay. Similar distinctions can be made between analytical, or assay, specificity and diagnostic specificity. Diagnostic specificity is affected by the assay’s analytic/ assay specificity as well as other factors including the assay’s analytical sensitivity in that a highly sensitive test may result in an increased probability of false-positives due to contamination [9].

For the remainder of the white paper, we will primarily focus on the following four terms to describe sensitivity and specificity: assay sensitivity (probability a sample correctly tests positive), diagnostic sensitivity (probability an individual is correctly identified as positive at a single testing event), assay specificity (probability a sample correctly tests negative), and diagnostic specificity (probability an individual is correctly identified as negative at a single testing event).

**Loss of sensitivity due to dilution.** When pooling specimen for group testing, it is essential to consider the dilution effects of pooling, i.e., the loss in diagnostic sensitivity due to diluting a positive sample with negative samples. We define dilution as the proportion of samples originally expected to test positive under individual testing (viral RNA > LLD) that are now expected to be in master pools in which the concentration of viral RNA is below the LLD. Pilcher et al. describe a scenario where a testing protocol that does not utilize pooling has a diagnostic sensitivity of 70% [7]. If pooling results in 10% dilution, the new diagnostic sensitivity will be  $70\% \times 90\% = 63\%$  (holding all variables other than pooling constant). In the R Shiny app and results, the diagnostic sensitivity under individual testing is assumed equal to assay sensitivity. Thus, in this paper, diagnostic sensitivity in a pooled testing scenario is equal to (assay sensitivity)  $\times$  (1-dilution).

Based on the viral dynamic model that the R Shiny app utilizes and the LLD of the SARS-CoV-2 assay, there is a 14-day window in which SARS-CoV-2 can be detected: that is, there are 14 days of infection in which SARS-CoV-2 concentration in nasopharyngeal specimen is at or above the LLD. After dilution, and assuming that there is at most one positive specimen per pool, the detection window will decrease as the master pool size increases.<sup>1</sup> The detection window and dilution effects described here and in the next section are specific to the assumed viral dynamic model and attributes of the SARS-CoV-2 assay, and will change as the viral dynamic assumptions or assay change.

The viral dynamics model itself is described in lab considerations.

**Maximum allowable dilution (MAD).** The maximum allowable dilution (MAD) is the proportion of the diagnostic sensitivity that one is willing to lose in order to pool together samples,

---

1. If there is more than one positive in a master pool, a specimen with a low concentration that would have fallen under the LLD after dilution might be "rescued" by the presence of another positive specimen in the same master pool. However, this situation is expected to be rare; the probability of multiple positives per pool is very small if pool size has been selected to optimize efficiency.

thus diluting positive samples with negative ones. The MAD value is determined *a priori*, and can range from 0 to 1.

**Maximum allowable pool size (MAPS).** For a given viral dynamics model, the MAD value implies a particular maximum allowable pool size (MAPS) – the largest value for the master pool size that will result in a diagnostic sensitivity loss no larger than the MAD.

The SARS-CoV-2 viral dynamics model for the R Shiny app assumes that there is a 14-day detection window. The equations that relate MAD and MAPS further assume that infected individuals present for testing uniformly during the detection window and that there is at most one positive specimen per master pool. The equations that relate MAD and MAPS are as follows [7,12]:

$$MAPS = 10^{(Rate\ of\ viral\ load\ increase \times detection\ window\ length \times MAD)}$$

$$MAD = \left[ \log_{10}(MAPS) \right] / (Rate\ of\ viral\ load\ increase \times detection\ window\ length)$$

Assuming the model of  $R = +1.0$  log<sub>10</sub> viral load per day (see *Viral dynamics model*) and a 14-day detection window, the equations are:

$$MAPS = 10^{(7 \times MAD)}$$

$$MAD = \left[ \log_{10}(MAPS) \right] / 7$$

In the following results, we have assumed that we set MAD a priori at 20% - we will not allow pooling to reduce diagnostic sensitivity more than 20% compared to individual testing. Another way to think of MAD = 0.20 is as a loss of 20% of the 14-day detection window, or 2.8 days (1.4 days on each side of the window). The MAPS for a maximum allowable dilution of 20% is calculated as:

$$MAPS = 10^{(7 \times 0.2)} \approx 25$$

Thus, in the following, we did not consider any master pools above size 25. Assuming the viral dynamics model is correct, this gives us a maximum dilution of 20%.

If people do not present uniformly in the detection window, the MAPS implied by a particular MAD could be different. Consider, for example, a community with routine, voluntary screening available for all individuals. Among infected persons with non-severe disease, those experiencing mild, non-specific symptoms (versus no symptoms) may be more likely to present for screening. Because viral load for SARS-CoV-2 is highest around the onset of symptoms (even if they are mild), infected persons in this testing population may disproportionately present around the peak viral load (days 5 through 10 of the 14-day window). In this scenario, one could actually have a higher MAPS for a given MAD than is expected under the assumption of uniform presentation. On the other hand, a screening program used exclusively among asymptomatic individuals in the general population, with no known exposures, would likely conform to the assumption that infected individuals present uniformly in the detection window.

**Positive predictive value (PPV).** Positive predictive value, simply, is the probability that given a positive test result, the sample or individual in question is truly positive. Pooling increases the PPV compared to individual testing because pooling strategies result in repeat testing of positive specimens, and thus an increase in the effective per-sample specificity, and thus the diagnostic specificity.

Specifically, in a minipooling (D2) approach with a truly negative master pool, a sample will only test positive if both (i) the master pool and (ii) the individual sample test (false) positive. In the absence of contamination and if the errors are independent, the probability of such an event for any particular sample is a function of assay specificity,  $(1 - Sp)^2$ . For a high specificity test (e.g., PCR often has specificity of 0.99 or greater) this means the probability that any

single sample will test false positive is  $(1-0.99)(1-0.99) = 0.0001$ , for an effective per-sample specificity of 0.9999.<sup>2</sup> Similar math shows that the effective per-sample specificity in D3 pooling could reach 0.999999. Since PPV is driven in large part by specificity (higher specificity leads to fewer false positives and higher PPV), this substantial increase in effective specificity of the assay under pooling can lead to large gains in PPV compared to individual testing.

However, if there is cross-contamination in the pooling process, the errors from round to round will not be independent and these considerations will not apply, nor will such calculations reflect reality. The risk of cross-contamination of samples during handling should be low if a consistent workflow and organizational scheme for pooling is implemented with good sterile technique, but the step of pooling samples introduces additional opportunities for cross-contamination to occur. Das et al. 2020 (doi: [10.1016/j.jcv.2020.104619](https://doi.org/10.1016/j.jcv.2020.104619)) provide a suggested workflow for minimization of such risks.

**Time to results.** Time to results is operationalized as the average number of rounds of testing required to obtain the final results, assuming that individual testing requires 1 round. For example, if at least 1 master pool tests positive in the D2 strategy, 2 rounds of testing will be required to obtain final results. If at least 1 master pool and sub-pool test positive in the D3 strategy, 3 rounds of testing will be required. Given at least one master pool tests positive, pooling will always increase the *minimum* time to results (when operationalized in this way) compared to individual testing, as individual testing can return results in a single round.

However, in practice, the amount of time required to obtain results for the population of interest may be substantially

---

2. The situation is of course more complicated, in that in an N:1 minipool, if the master pool tests positive, there are N opportunities for a second false positive. Likewise, if the pool contains a true positive sample which is correctly identified, then there are N-1 opportunities for a false positive at the original assay specificity. Additional biostatistical details for this and related considerations can be found in [4].



less with pooling versus individual testing due to the reduced number of assays run in the context of pooling [7,13].

## BIOSTATISTICS OF POOLING

As noted elsewhere, the biostatistics of pooling are complex. For details, we suggest starting with:

- Bilder C, Iwen P, Abdalhamid B, Tebbs J, and McMahan C. (2020). Tests in short supply? Try group testing. *Significance* 17(3), 15-16
- Bilder C. (2019). Group testing for estimation. Wiley StatsRef: Statistics Reference Online. <https://doi.org/10.1002/9781118445112.stat08231>
- Bilder C. (2019). Group testing for identification. Wiley StatsRef: Statistics Reference Online. <https://doi.org/10.1002/9781118445112.stat08227>

and then additional, more technical literature including:

- Kim HY, Hudgens MG, Dreyfuss JM, Westreich DJ, Pilcher CD. Comparison of group testing algorithms for case identification in the presence of test error. *Biometrics*. 2007 Dec;63(4):1152-63. doi: 10.1111/j.1541-0420.2007.00817.x. PMID: 17501946
- Litvak E, Tu XM and Pagano M, 1994. Screening for the presence of a disease by pooling sera samples. *Journal of the American Statistical Association*, 89(426), pp.424-434.
- Johnson NL, Kotz S, and Wu X. (1991). *Inspection Errors for Attributes in Quality Control*. New York: Chapman and Hall Ltd.

## HOW THE CALCULATOR WORKS/ OPTIMIZES POOL SIZE

The R Shiny app can be found by visiting [gillingscovid19.unc.edu](https://gillingscovid19.unc.edu) and searching *pooling*.

Additionally, the latest version of the SARS-CoV-2 web calculator, which was first made available in August 2020 but has been updated, can be found here:

<http://www.bios.unc.edu/~mhudgens/SARS-CoV-2.pooling.home.html>

The results from the R Shiny app and web calculator are identical. When using the SARS-CoV-2 R Shiny app, there are 5 inputs that the user must specify: assay sensitivity (defaults to 0.95); maximum allowable dilution (MAD) (defaults to 0.2); assay specificity (defaults to 0.99); prevalence (or range of prevalence values); pooling strategy of interest (D2, D3, and/or A2m). The maximum allowable pool size (MAPS) is automatically calculated for the user using the equation described in the previous section (see *Definitions/explanation of terms*) and the input MAD value. When using the web calculator, the user must specify the MAPS (though clicking on the “compute” button will populate this value if MAD has been specified), and results for all three pooling strategies are automatically presented.

The R Shiny app (using R code) and the web calculator (using HTML code) calculate the optimal pool size for three pooling designs: D2, D3, and the A2m (though we do not present results for A2m pooling below, as it is likely too complicated, with relatively little advantage over D3 pooling, for labs to pursue in general). The calculator reports the optimal pool size for each design, and the efficiency and PPV associated with that pool size.

For each pooling design, the optimal pool size is the pool size with the lowest value for efficiency (i.e., most efficient). For the D2 design, given

the 5 inputs specified, the program calculates the efficiency for every potential master pool size in the range starting at the value of MAPS and ending at 2 (because 1 is simply individual testing). For example, if one decides that the maximum allowable dilution is 0.2, then one would specify a MAPS of 25 (or 25 would be auto-calculated using the R Shiny app), and the program would calculate efficiency for every integer between 2 and 25. The pool size between 2 and 25 with the best value for efficiency would be selected as the optimal pool size.

For the D3 and A2m designs, the program calculates efficiency for every integer between 2 and the MAPS value *that is also a square number*. For example, given MAPS=25, the program would calculate efficiency for four potential master pool sizes (4, 9, 16, and 25), and the pool size with the lowest value for efficiency would be reported as the optimal.

After the optimal pool size is selected, the PPV is calculated based on a function of optimal pool size, the updated diagnostic sensitivity, assay specificity, and prevalence.

## UPDATES TO THE R SHINY APP/WEB CALCULATOR SINCE AUGUST 2020

A version of the web calculator was made available in August 2020. Since August 2020, there have been two major changes to the code used to optimize pool size; these changes are reflected in the R Shiny app and latest version of the published web calculator. First, diagnostic sensitivity is now updated as the calculator optimizes pool size, and second the final, diagnostic sensitivity is updated to reflect the true dilution expected to be realized in the final, optimal pool.

Previously, the efficiency and PPV were calculated using the diagnostic sensitivity associated with the MAPS (i.e.,  $(1-\text{MAD}) \times \text{assay sensitivity}$ ), regardless of the pool size for which efficiency and PPV were being

calculated. As of December 2020, all code calculates efficiency for a given pool size using the diagnostic sensitivity updated for each current pool size, and calculates PPV for the optimal pool size using the diagnostic sensitivity expected under that optimal pool size.

## **V. LAB CONSIDERATIONS FOR POOLING COVID**

SARS-CoV-2 RT-PCR Tests. Real-time reverse transcriptase polymerase chain reaction (RT-PCR) tests are a type of nucleic acid amplification test (NAAT) used to detect the presence of specific sequences of RNA. After extracting RNA from a sample, RT-PCR transcribes RNA into DNA, the DNA is amplified, and as the DNA accumulates a fluorescent signal emerges indicating its presence. The number of cycles of amplification needed to produce the fluorescent signal is termed the cycle threshold (Ct) – a measure that is inversely correlated with the concentration of RNA in the sample. RT-PCR is highly sensitive (high analytical sensitivity) and frequently used for virus detection, though it is costly and resource-intensive to implement.

The Centers for Disease Control and Prevention (CDC) have developed two real-time RT-PCR tests for the detection of SARS-CoV-2 RNA: the 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel and the CDC Influenza SARS-CoV-2 (Flu SC2) Multiplex Assay. These tests are approved by the U.S. Food and Drug Administration (FDA) under an emergency use authorization (EUA). Both tests are designed to identify SARS-CoV-2 RNA from upper (e.g., nasopharyngeal) or lower (e.g., sputum) respiratory samples.

# EMERGENCY USE AUTHORIZATION FOR SARS-COV-2 POOLED TESTING

The “Quest SARS-CoV-2 rRT-PCR” was authorized for use by the FDA under an EUA, and is designed to detect SARS-CoV-2 RNA in upper or lower respiratory samples as well. This EUA permits individual testing with the assay, as well as pooled specimen testing with master pools of size 4 or less. This assay has an LLD of 136 copies/mL – at this concentration, the probability of correctly classifying a specimen as positive is 95%. A Ct value of <40 is considered a positive result. The specificity (i.e., assay specificity) reported in the package insert for this assay is 100%.

For directions related to the use of SARS-CoV-2 RT-PCR tests, information about required supplies, and further details of the assay characteristics, please refer to the citations referenced.<sup>3</sup>

## LABORATORY CONSIDERATIONS

The implementation of pooling strategies tends to increase the complexity of laboratory procedures for processing and analyzing samples relative to individual testing of individual samples. This added complexity can be addressed through the use of pooling robots, which have the potential to save time and prevent errors, but require additional resources for their purchase, setup, and programming. The complexity added to a testing process by implementing pooled testing depends on the strategy being used, with 2- or 3-stage hierarchical pooling likely presenting a lesser organizational and analytic challenge than array pooling, and combinatorial pooling requiring either a pooling robot or a great deal of organizational effort and risk of error.

---

3. CDC tests: <https://www.fda.gov/media/134922/download> and <https://www.fda.gov/media/139743/download>; Quest <https://www.fda.gov/media/136231/download>

# VIRAL DYNAMICS MODEL

The SARS-CoV-2 web calculator and R Shiny app assume a model of the typical viral load trajectory for an individual who is infected with SARS-CoV-2 and symptomatic, but not-critically ill. The calculator assumes a 14-day detection window, in which viral RNA increases steadily at a rate of  $+1.0 \log_{10}$  viral load per day for 4 days, plateaus at a peak of  $4.2 \log_{10}$  viral load for 6 days, and steadily decreases at a rate of  $-1.0 \log_{10}$  viral load per day for 4 days – Figure 4 (left). A sample that is collected at the beginning or end of the detection window (compared to the peak) is more likely to be undetected by a pooled testing algorithm, because it is closer to the LLD and thus more likely to be in a master pool in which the viral RNA concentration is below the LLD (unless it is “rescued” by the presence of another positive sample in the same master pool) – Figure 4 (below) [8]. Finally, it is assumed that cases present for testing uniformly during the detection window. Thus, any question related to the proportion of positive samples “lost” due to the decreased diagnostic sensitivity of pooling can be framed as the proportion of days during the detection window that are no longer detectable due to dilution.

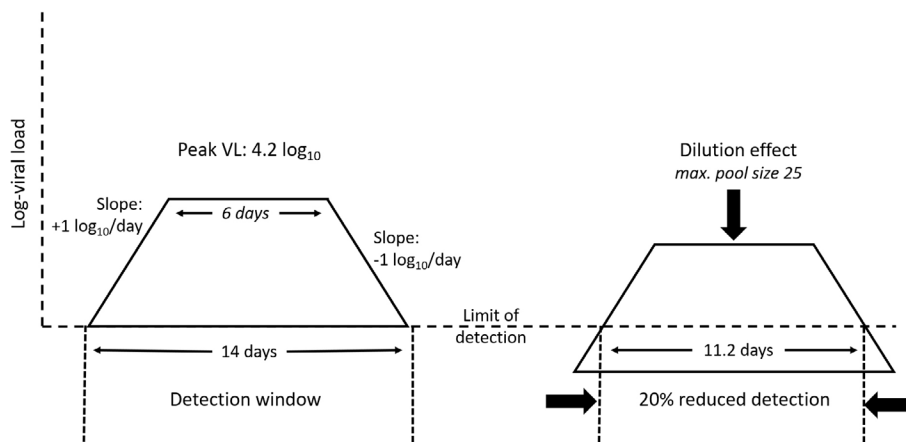


Figure 4: Model for nasopharyngeal (NP) RNA in acute severe acute respiratory syndrome–coronavirus 2 (SARS CoV-2) infection. From (Pilcher, Westreich, Hudgens. Journal of Infectious Diseases. 2020).

Pilcher et al. confirmed that the assumed viral load dynamic model produces a distribution of viral loads that is in accordance with viral load results from recent clinical studies of non-critical COVID-19 cases [7]. The assumed viral load dynamics are based on literature review [14–20] and are subject to change as more information becomes available.

For example, there is literature that supports the possibility of a more asymmetric viral load trajectory: a rapid increase in viral load after infection, a peak in viral load around the time of symptom onset, followed by a slower decline in viral load with low, but detectable, viral shedding past day 14 of infection [21–25]. Cevik et al. conducted a meta-analysis of 79 SARS-CoV-2-related studies and found that the longest period of viral shedding reported was 83 days in the upper respiratory tract for one patient [21]. While it is possible that some infected individuals have an extended period of viral shedding during convalescence, it may not correspond with infectiousness (viral load may be very low, and cultivable virus may rarely persist beyond 10 days post symptom onset for mild/moderate cases); thus, these may not be cases that are highly important to detect [21,22,24,26,27]. The assumption that persons with asymptomatic infection have similar viral load trajectories as those with mild/moderate disease seems to be supported at this time, though the viral load kinetics may be different for persons with severe disease [19,21,26,28].

If any updates to the viral load dynamic model occur in the future, a log of these changes will be recorded and available on the calculator website. We note that if it is in fact the case that viral load declines more slowly, and extends beyond day 14 of infection, then (broadly) we anticipate that the calculations here related to sensitivity loss are conservative, and that true sensitivity loss would be smaller than we anticipate in our model

(or, alternatively, that larger pool sizes may be acceptable while still limiting loss of diagnostic sensitivity due to dilution).

## VI. PREDICTED RESULTS OF POOLING FOR SARS-COV-2

Table 1 presents results for the SARS-CoV-2 pooling as reported in Pilcher, Westreich, Hudgens[7]<sup>4</sup> with a range of prevalence values from 0.001 to 0.10 under the default settings of assay sensitivity=0.95, MAD=0.2 (and thus MAPS=25), and assay specificity=0.99. For prevalence values ranging from 0.001 to 0.10, the D2 efficiency ranges from 0.0685 to 0.5552 and the D3 efficiency ranges from 0.0497 to 0.4966. Within each pooled testing strategy, pooling is most efficient at the lowest prevalence. As prevalence increases, pooled testing for a given strategy becomes less efficient, but within the range examined here pooled testing remains more efficient than individual testing. For example, even at prevalence 10%, the D3 strategy can obtain results for 2,014 individuals per 1,000 tests (versus 1,000 results obtained under individual testing).<sup>5</sup> Furthermore, pooling still remains more efficient than individual testing at prevalence values of 11% to 20% (see Figure 6 – graphical output from the R Shiny app for the D2 and D3 strategies).

Within the D2 pooling strategy, optimal pool size decreases, or remains constant, as prevalence increases (Table 1, Figure 5a). Optimal pool size decreases from 25 at prevalence=0.1% to 4 at prevalence=10%, and remains 4 at prevalence=15% or 20%. As prevalence increases, large master pools have an increased probability of containing a positive result. In order to retain the benefit of pooling at higher prevalence values, a smaller master pool size is typically favored to

---

4. Note that the web calculator may produce slightly different results, because the publication relied on additional code to confirm and sometimes correct findings from the web calculator. However the results from the web calculator were never wrong by much and can be relied upon broadly for planning purposes. Inconsistencies will be resolved in future work.

5. Though it may not be worth additional laboratory effort to pool in such a high prevalence setting. In addition, we obtain this number by taking the reciprocal of the efficiency (to more decimal places than is shown in Table 1) and then multiplying by 1000.



ensure that some master pools test negative. However, at higher prevalence values, when master pools are extremely unlikely to be negative, larger pool sizes may once again become more favorable. In the D3 strategy, the optimal master pool size drops from 25 to 16 at prevalence=5%, and then returns to 25 at prevalence=10%. Similarly, optimal pool size returns to 25 for the D2 strategy at prevalence=30%.

At the prevalence levels examined in Table 1, the more complex strategy (D3 versus D2, see also Figure 5b) is more efficient at a given prevalence. Per 1,000 tests at prevalence=1%, D3 pooling can obtain results for ~8,800 individuals versus ~5,500 individuals for D2 pooling. However, as prevalence increases toward 100% and fewer of the pools test negative, the more complex D3 strategy has similar efficiency to D2 pooling, and ultimately individual testing.

Both pooling strategies, at all prevalence levels explored, result in large gains in positive predictive value (which, again, assumes independence of errors between rounds of testing, including no contamination). At each prevalence level, PPV is higher in the pooled scenario versus individual testing, with the absolute difference between the pooled PPV and individual PPV largest at lowest prevalence. This is because individual testing is particularly poor at low prevalence levels: at prevalence=0.1%, only 9% of individuals who test positive are expected to be truly positive. In the D2 strategy at prevalence=0.1%, pooled PPV is 73% and individual PPV is 9%. At prevalence=10%, pooled PPV for D2 is 98% and individual PPV is 91%. At prevalence=20%, pooled PPV is 98% for D2, compared to 96% for individual testing. At each prevalence level <10%, PPV is higher for D3 versus D2; at prevalence levels of 10%-20%, the PPV for D3 and D2 are comparable.

As expected, the average time to results for both pooling strategies at every prevalence level examined is greater than the average time to results for individual testing, which is always equal to one round of testing given our operationalization of this outcome. The average

time to results under pooling increases as prevalence increases. For example, in the D3 strategy, average time to results is 1.03 rounds at prevalence=0.1% and increases to 2.02 rounds at prevalence=10%. At each prevalence level, the D3 algorithm has a higher average time to results than the D2 algorithm, which was expected given D3 testing has one additional possible round of testing compared to D2 testing.

On time to results, however, it is important to note that if a lab had more-limited capacity for throughput, pooling could provide substantial advantages in time-to-results. E.g., if a particular lab had the capacity to run a total of one round of 100 individual tests per day, and 1000 specimens arrive at the lab, then turn around time for those 1000 specimens would be a mean of 5.5 days to turnaround for each specimen, and it would take 10 days to time to results for all individuals. On the other hand, at a prevalence of 1%, 10:1 minipools might take one day to construct 100 pools of 10 specimens each; one day to process the 100 pools; and one day to repeat test specimens in any positive pools (of which we expect approximately 10, comprising 100 more individual samples). This would require 3 days to results for all individuals, or a mean less than 3 days. This might prove a substantial advantage if, for example, positive results are being used for contact tracing.

Even though diagnostic sensitivity is reduced by pooling, both pooling strategies at every prevalence level examined identify more true positives per 1,000 tests compared to individual testing. At prevalence=10%, individual testing is expected to identify 95 true positives, whereas D2 and D3 pooling is expected to identify 156 and 153 true positive cases, respectively. Unsurprisingly, number of true positives identified increases as prevalence of disease increases. At each prevalence level <10% examined, the D3 strategy identified the highest number of true positives per 1,000 tests, whereas at prevalence levels 10-20%, D2 identified slightly more true positives than D3 per 1,000 tests.

## VII. PRACTICAL GUIDANCE ON POOLING IMPLEMENTATION

**Sorting specimens by clinical presentation (symptomatic vs. asymptomatic) prior to pooling.** Analysis is currently underway to assess the extent to which screening individuals with a low probability of testing positive (asymptomatic, no known exposures) to pool separately from symptomatic individuals could improve pooling efficiency. By lowering the effective prevalence in the asymptomatic pools, larger and more efficient pool sizes may be possible. However, the extent to which this additional screening is worthwhile necessarily depends on the effort involved in screening and whether the screening strategy can effectively differentiate lower risk individuals.

**Considerations for when and how to update the input for prevalence in the R Shiny app.** It may be logistically impractical to frequently update the pooling strategy implemented in a lab based on changing prevalence in the area, and changing the pooling algorithm would likely be reserved for substantial shifts in the reported prevalence. If adjustments are to be made to the pooling strategy in light of a change in prevalence, they should be made based on the percent test positivity in the region, rather than population prevalence. Averaged estimates of test positivity over more than one day (versus test positivity on a single day) are also likely to be more stable indicators on which to base optimal pooling strategies.

**Antigen testing and the importance of screening frequency for case detection.** Throughout this paper, we have focused on diagnostic sensitivity, defined as the proportion of cases correctly identified as positive *at a single time point*, which is a product of assay sensitivity and dilution due to pooling in our calculations. However, the overall goal of screening for case detection is to identify positive individuals *at some point during their infection*, ideally near the beginning of

infection to prevent transmission to others. In a recent article, Mina et al. emphasize the importance of considering the sensitivity of an overall testing approach, and theoretically compare the probability that an infected individual will be identified as positive at some point during their infection using occasional PCR testing versus more routine antigen testing – see Mina et al., Figure (not numbered) [29].

Compared to antigen testing, PCR testing can detect a smaller amount of virus in specimen and has higher diagnostic sensitivity accordingly, but is more resource-intensive and costly to implement, and may have longer turn-around time to results [29]. Thus, antigen testing, which may miss more cases than PCR during a single testing event, may actually identify more positive cases over a period of time due to the higher frequency with which the antigen test can be administered [29]. Importantly, this elevated test frequency along with the quicker turn-around time of results may facilitate identification of individuals before or during their peak viral load when methods to prevent transmission are most needed [29]. As well, antigen testing may be less likely to pick up individuals with lower viral loads, who are less likely to transmit the virus to others. Pooled PCR testing may be a strategy analogous to antigen testing: compared to individual PCR testing, pooled PCR testing saves money and resources at the expense of lower diagnostic sensitivity. *Pooled PCR testing may likewise be a mechanism to increase the frequency of screening in populations and the overall detection of positive individuals at some point during their infection.* However, whether pooled PCR testing has the potential to decrease the turn-around time of results (given the logistical challenge it may impose on a lab) needs to be carefully considered.

**Application of pooling to antibody testing.** The methods and results presented in this paper focus on pooled PCR testing for the detection of virus. However, antibody testing will play an essential role in further understanding the COVID-19 pandemic (e.g., understanding total number of COVID-19 cases and the proportion

who experience asymptomatic versus symptomatic infection), and demand for widespread antibody testing may make pooled antibody testing desirable. However, the variety of serological assays for SARS-CoV-2 antibodies, as well as a lack of current understanding of SARS-CoV-2 antibody kinetics and how they vary among individuals and over time, make the direct application of the methods presented here and accompanying R Shiny app questionable at best. The extent to which underlying assumptions about the impact of pooling on diagnostic sensitivity and specificity are reasonable would depend on the characteristics of the assay being used; the variety of assay approaches, antibody targets, and cutoff values used would make a generalized commentary on pooling of samples for this purpose inadequate or misleading.

Test performance attributes of FDA approved serology assays can be found here: <https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance>

## VIII. SUMMARY AND TAKEAWAYS

Here we have shown the theoretical tradeoffs in the total number of tests required for case detection, the number of rounds of testing required to get results, and the positive predictive value of the results that can emerge from pooling. However, whether pooling is “worth it,” overall – and potential for delay and error involved in implementing the optimal pooling algorithm – depend on the context. This R Shiny app and document provide guidance on selecting optimal pooling strategies, however the ultimate decision of whether or not to pool may be complex and driven by more than the 5 quantitative inputs required for the Shiny app to run. Furthermore, depending on context, users may want to optimize pooling for an outcome other than efficiency, such

as number of results returned within 3 days, as we discussed above.

When using pooling methods, the optimal testing algorithm and pool size depend not only on the underlying sensitivity and specificity of the assay used, but the proportion of samples that are truly positive, which means that the volatility of test-positivity rates in an ongoing epidemic may require recalibration to maintain the efficiency and positive predictive value of pooled testing over time. This may be achieved through periodic prevalence estimation through broader pooled testing, reported test positivity rates in an area, or cut points that dictate the pool size of testing moving forward based on the findings of the previous period.

There are several limitations to consider when using the R Shiny app to implement the pooling strategies described within this paper. First, all of the input probabilities presented (e.g., assay sensitivity = 95%) are not deterministic, i.e., there is variability in each probability and the input chosen is only an estimate. Thus, while we present a single estimate for efficiency for a given set of parameters, there is in fact variability around the estimate. Second, the calculations encoded by the R Shiny app/web calculator generally make the assumption that there is at most one positive specimen per pool. For instance, the impact of pool size on dilution assumes that there is at most one positive per pool. If however, there is more than a single positive specimen in a positive pool, the dilution due to pooling will be less extreme than the results presented. Third, our definition of diagnostic sensitivity is only affected by assay sensitivity and dilution due to pooling. In practice, a positive individual could ultimately be misclassified as negative/indeterminate during a single testing event for other reasons, such as inadequate nasopharyngeal swabbing or a data entry error. However, individual testing is also susceptible to these errors.

In testing applications where individuals with higher probability of testing positive are being screened (>1% prevalence), we feel that

simpler 2- and 3-stage hierarchical pooling are generally preferable to more complex options like array pooling (such as A2m pools) or combinatorial pooling strategies. In the context of SARS-CoV-2, the more complex pooling strategies are generally being proposed and implemented to detect asymptomatic individuals with a low prior probability of infection. In general, higher background positivity rates in the samples being tested lend themselves to simpler pooling strategies and smaller pooling sizes.

The included guidance shows a clear pathway forward if state laboratories wish to implement simple pooling algorithms for more efficient detection of samples, which may bear fruit even at current (December 2020) high rates of positivity. Laboratories wishing to implement such pooling can feel free to get in contact with the pooling lead at Gillings Center for Coronavirus Testing, Screening, and Surveillance, Dr. Daniel Westreich ([djw@unc.edu](mailto:djw@unc.edu)).

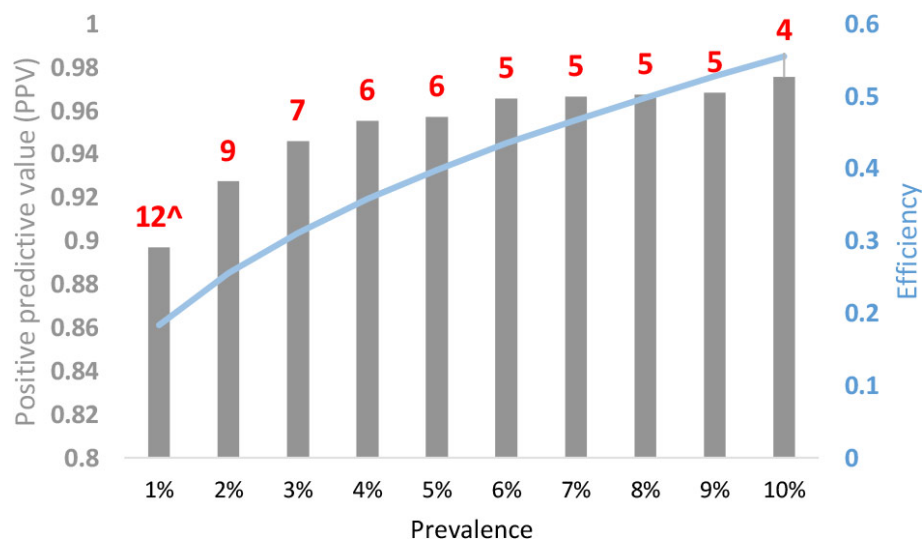
**Table 1.** Testing characteristics for two pooling strategies (D2, D3)\*

	Scenario				Performance relative to individual testing				
Pooling strategy	Prevalence	Suggested algorithm	Time to results	Efficiency	Sens. loss, pooled v. individual testing	Positive predictive value		# true positives ID'd per 1,000 tests	
						Pooled	Individual	Pooled	Individual
D2	0.001	25:1	1.03	0.07	0.20	0.73	0.09	11.1	0.95
	0.005	18:1	1.08	0.13	0.18	0.84	0.32	29.5	4.75
	0.01	12:1	1.10	0.18	0.15	0.90	0.49	43.8	9.5
	0.05	6:1	1.23	0.40	0.11	0.96	0.83	106.2	47.5
	0.1	4:1	1.31	0.56	0.09	0.98	0.91	156.4	95
D3	0.001	25:5:1	1.03	0.05	0.20	0.96	0.09	15.3	0.95
	0.005	25:5:1	1.12	0.08	0.20	0.96	0.32	47.9	4.75
	0.01	25:5:1	1.22	0.11	0.20	0.96	0.49	66.7	9.5
	0.05	16:4:1	1.57	0.32	0.17	0.97	0.83	121.9	47.5
	0.1	25:5:1	2.02	0.50	0.20	0.97	0.91	153.1	95

\*Adapted from (Pilcher et al., 2020).



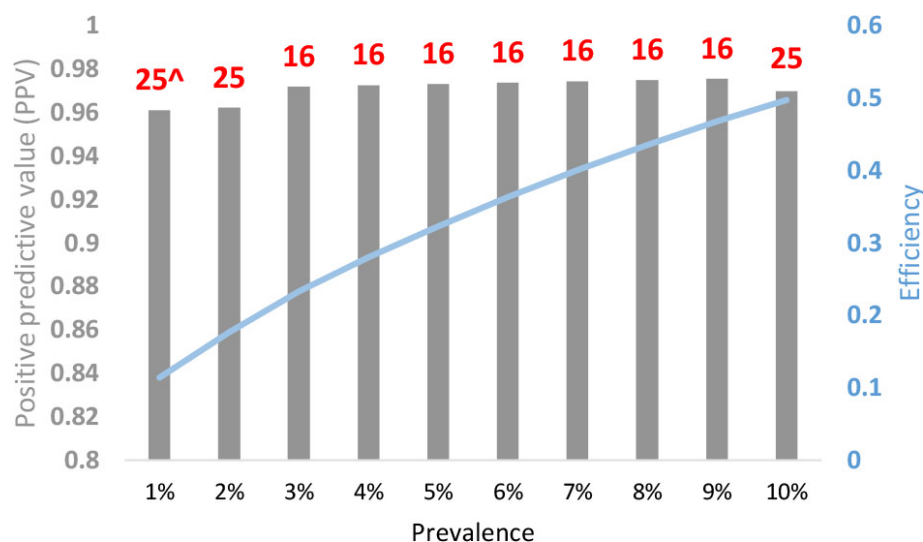
**Figure 5a.** D2 pooling strategy – efficiency and optimal pool size over a range of prevalence values (1% to 10%)\*



\*Calculator settings: assay sensitivity = 0.95; maximum allowable dilution = 0.20; maximum allowable pool size =25; assay specificity = 0.99

▲Red value indicates the optimal master pool size

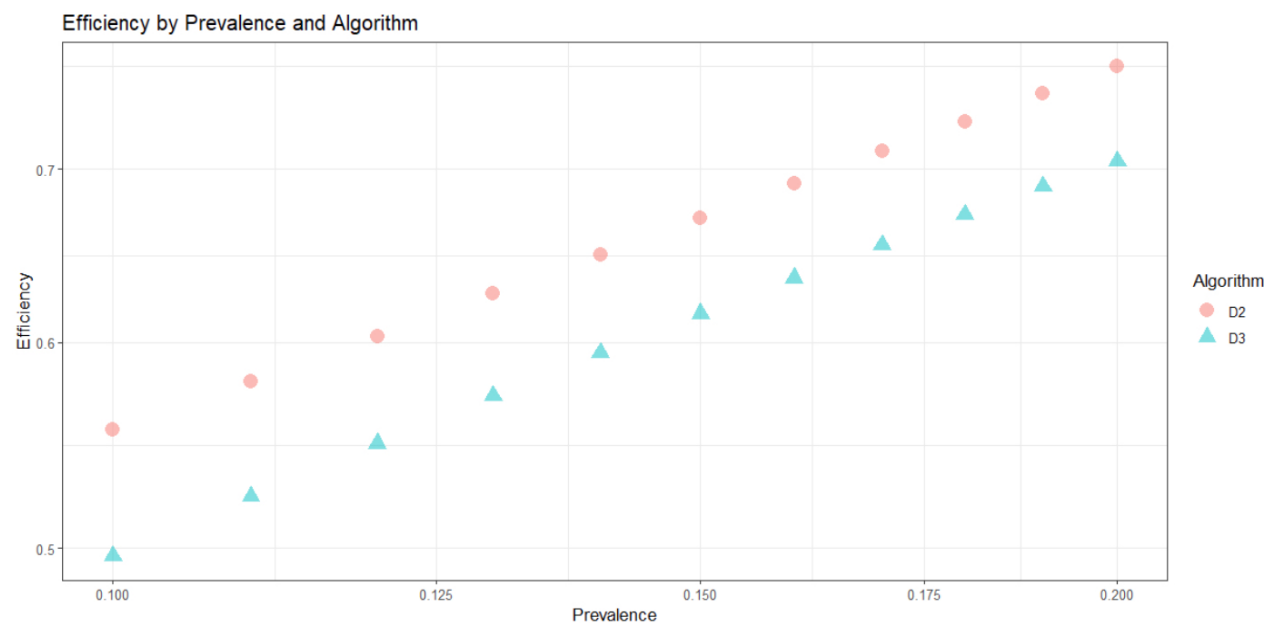
**Figure 5b.** D3 pooling strategy – efficiency and optimal pool size over a range of prevalence values (1% to 10%)\*



\*Calculator settings: assay sensitivity = 0.95; maximum allowable dilution = 0.20; maximum allowable pool size =25; assay specificity = 0.99

▲Red value indicates the optimal master pool size

**Figure 6.** Graphical output from the R Shiny app for the D2 and D3 pooling strategy – efficiency over a range of prevalence values (10% to 20%). (*Calculator settings: assay sensitivity = 0.95; maximum allowable dilution = 0.20; maximum allowable pool size = 25; assay specificity = 0.99*)



## CITATIONS

1. Dorfman R. The Detection of Defective Members of Large Populations. *Ann Math Stat Published Online First*: 1943. doi:10.1214/aoms/1177731363
2. Zenios SA, Wein LM. Pooled testing for HIV prevalence estimation: Exploiting the dilution effect. *Stat. Med.* 1998. doi:10.1002/(SICI)1097-0258(19980715)17:13<1447::AID-SIM862>3.0.CO;2-K
3. Finucan HM. The Blood Testing Problem. *Appl Stat Published Online First*: 1964. doi:10.2307/2985222
4. Kim HY, Hudgens MG, Dreyfuss JM, Westreich DJ, Pilcher CD. Comparison of group testing algorithms for case identification in the presence of test error. *Biometrics Published Online First*: 2007. doi:10.1111/j.1541-0420.2007.00817.x
5. Shental N, Levy S, Skorniakov S, Wuvshet V, Shemer-Avni Y, Porgador A, et al. Efficient high throughput SARS-CoV-2 testing to detect asymptomatic carriers. *medRxiv* 2020; :2020.04.14.20064618.
6. Mutesa L, Ndishimye P, Butera Y, Souopgui J, Uwineza A, Rutayisire R, et al. A strategy for finding people infected with SARS-CoV-2: optimizing pooled testing at low prevalence. *medRxiv* 2020; :2020.05.02.20087924.
7. Pilcher CD, Westreich D, Hudgens MG. Group Testing for Sars-Cov-2 to Enable Rapid Scale-Up of Testing and Real-Time Surveillance of Incidence. *J Infect Dis Published Online First*: 2020. doi:10.1093/infdis/jiaa378
8. Cleary B, Hay JA, Blumenstiel B, Gabriel S, Regev A, Mina MJ. Efficient prevalence estimation and infected sample identification with group testing for SARS-CoV-2. ; 2020. doi:10.1101/2020.05.01.20086801
9. Saah AJ, Hoover DR. "Sensitivity" and "specificity" reconsidered: The meaning of these terms in analytical and diagnostic settings. *Ann. Intern. Med.* 1997. doi:10.7326/0003-4819-126-1-199701010-00026
10. Forootan A, Sjöback R, Björkman J, Sjögreen B, Linz L, Kubista M. Methods to determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR). *Biomol Detect Quantif Published Online First*: 2017. doi:10.1016/j.bdq.2017.04.001
11. OPTI SARS-CoV-2 RT-PCR Test. <https://www.fda.gov/media/137739/download>

12. Westreich DJ, Hudgens MG, Fiscus SA, Pilcher CD. Optimizing screening for acute human immunodeficiency virus infection with pooled nucleic acid amplification tests. *J Clin Microbiol* Published Online First: 2008. doi:10.1128/JCM.00787-07
13. Agoti CN, Mutunga M, Lambisia AW, Kimani D, Cheruiyot R, Kiyuka P, et al. Pooled testing conserves SARS-CoV-2 laboratory resources and improves test turn-around time: experience on the Kenyan Coast. *Wellcome Open Res* Published Online First: 2020. doi:10.12688/wellcomeopenres.16113.1
14. Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. *Clin Infect Dis* Published Online First: 2020. doi:10.1093/cid/ciaa344
15. Tan W, Lu Y, Zhang J, Wang J, Dan Y, Tan Z, et al. Viral Kinetics and Antibody Responses in Patients with COVID-19. *medRxiv* 2020; :2020.03.24.20042382.
16. Pan Y, Zhang D, Yang P, Poon LLM, Wang Q. Viral load of SARS-CoV-2 in clinical samples. *Lancet Infect. Dis.* 2020. doi:10.1016/S1473-3099(20)30113-4
17. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* Published Online First: 2020. doi:10.1038/s41586-020-2196-x
18. Xu T, Chen C, Zhu Z, Cui M, Chen C, Dai H, et al. Clinical features and dynamics of viral load in imported and non-imported patients with COVID-19. *Int J Infect Dis* Published Online First: 2020. doi:10.1016/j.ijid.2020.03.022
19. Liu Y, Yan LM, Wan L, Xiang TX, Le A, Liu JM, et al. Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect. Dis.* 2020. doi:10.1016/S1473-3099(20)30232-2
20. To KKW, Tsang OTY, Leung WS, Tam AR, Wu TC, Lung DC, 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. *Lancet Infect Dis* Published Online First: 2020. doi:10.1016/S1473-3099(20)30196-1
21. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV-1 and MERS-CoV Viral Load Dynamics, Duration of Viral Shedding and Infectiousness: A Living Systematic Review and Meta-Analysis. *SSRN Electron J* Published Online First: 2020. doi:10.2139/ssrn.3677918
22. Meyerowitz EA, Richterman A, Gandhi RT, Sax PE. Transmission of SARS-CoV-2: A Review of Viral, Host, and Environmental Factors. *Ann Intern Med* Published Online First: 2020. doi:10.7326/m20-5008

23. Xiao AT, Tong YX, Zhang S. Profile of RT-PCR for SARS-CoV-2: a preliminary study from 56 COVID-19 patients. Clin Infect Dis Published Online First: 2020. doi:10.1093/cid/ciaa460
24. Young BE, Ong SWX, Ng LFP, Anderson DE, Chia WN, Chia PY, et al. Viral Dynamics and Immune Correlates of Coronavirus Disease 2019 (COVID-19) Severity. Clin Infect Dis Published Online First: 2020. doi:10.1093/cid/ciaa1280
25. Long QX, Tang XJ, Shi QL, Li Q, Deng HJ, Yuan J, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. Nat Med Published Online First: 2020. doi:10.1038/s41591-020-0965-6
26. Singanayagam A, Patel M, Charlett A, Bernal JL, Saliba V, Ellis J, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. Eurosurveillance Published Online First: 2020. doi:10.2807/1560-7917.ES.2020.25.32.2001483
27. Walsh KA, Jordan K, Clyne B, Rohde D, Drummond L, Byrne P, et al. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. J. Infect. 2020. doi:10.1016/j.jinf.2020.06.067
28. Lee S, Kim T, Lee E, Lee C, Kim H, Rhee H, et al. Clinical Course and Molecular Viral Shedding among Asymptomatic and Symptomatic Patients with SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea. JAMA Intern Med Published Online First: 2020. doi:10.1001/jamainternmed.2020.3862
29. Mina MJ, Parker R, Larremore DB. Rethinking Covid-19 Test Sensitivity – A Strategy for Containment. N Engl J Med Published Online First: 2020. doi:10.1056/nejmp2025631
30. Klasse PJ. Neutralization of Virus Infectivity by Antibodies: Old Problems in New Perspectives. Adv Biol 2014; 2014:157895.
31. Baumgarth N, Nikolich-Zugich J, Lee FE-H, Bhattacharya D. Antibody Responses to SARS-CoV-2: Let's Stick to Known Knowns. J Immunol Published Online First: 2020. doi:10.4049/jimmunol.2000839
32. Ibarrondo FJ, Fulcher JA, Goodman-Meza D, Elliott J, Hofmann C, Hausner MA, et al. Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-19. N Engl J Med Published Online First: 2020. doi:10.1056/nejmc2025179
33. Ripberger TJ, Uhrlaub JL, Watanabe M, Wong R, Castaneda Y, Pizzato HA, et al. Orthogonal SARS-CoV-2 Serological Assays Enable Surveillance of Low-Prevalence Communities and Reveal Durable Humoral Immunity. Immunity Published Online First: 2020. doi:10.1016/j.immuni.2020.10.004

34. Isho B, Abe KT, Zuo M, Jamal AJ, Rathod B, Wang JH, et al. Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Sci Immunol* Published Online First: 2020. doi:10.1126/sciimmunol.abe5511
35. Dan JM, Mateus J, Kato Y, Hastie KM, Faliti CE, Ramirez SI, et al. Immunological memory to SARS-CoV-2 assessed for greater than six months after infection. *bioRxiv* 2020.
36. Lin Q, Zhu L, Ni Z, Meng H, You L. Duration of serum neutralizing antibodies for SARS-CoV-2: Lessons from SARS-CoV infection. *J. Microbiol. Immunol. Infect.* 2020. doi:10.1016/j.jmii.2020.03.015