

Evaluation of Group Testing for SARS-CoV-2 RNA

Nasa Sinnott-Armstrong* Daniel L. Klein† Brendan Hickey‡

March 25, 2020

Abstract

During the current COVID-19 pandemic, testing kit and RNA extraction kit availability has become a major limiting factor in the ability to determine patient disease status and accurately quantify prevalence. Current testing strategies rely on individual tests of cases matching restrictive diagnostic criteria to detect SARS-CoV-2 RNA, limiting testing of asymptomatic and mild cases. Testing these individuals is one effective way to understand and reduce the spread of COVID-19.

Here, we develop a pooled testing strategy to identify these low-risk individuals. Drawing on the well-studied group testing literature, modeling suggests practical changes to testing protocols which can reduce test costs and stretch a limited test kit supply. When most tests are negative, pooling reduces the total number of tests up to four-fold at 2% prevalence and eight-fold at 0.5% prevalence. At current SARS-CoV-2 prevalence, randomized group testing optimized per country could double the number of tested individuals from 1.85M to 3.7M using only 671k more tests.

This strategy is well-suited to supplement testing for asymptomatic and mild cases who would otherwise go untested, and enable them to adopt behavioral changes to slow the spread of COVID-19.

Introduction

Group testing was first described in [5] as a technique for screening United States Army recruits for syphilis. In a common extension, samples are organized into a matrix, then each row and column of the matrix are tested [13, 6]. Under either approach, it is well understood that given a sensitive test and low disease prevalence, testing capacity increases dramatically.

We propose adapting this technique for diagnosis of SARS-CoV-2, where COVID-19 prevalence is low and testing is limited. This is in part motivated by work suggesting that population-scale testing for SARS-CoV-2 might effectively reduce asymptomatic transmission [1]. During the past week, as we finalized this manuscript, its application to SARS-CoV-2 testing has been widely suggested and demonstrated [20, 19, 8].

Since group testing is most efficient when prevalence is low, we propose using it to test individuals who do not meet current test criteria. This obviates many concerns around group testing efficacy on the individual level, as critical care resources are not essential for asymptomatic cases. Individuals with a high case probability can be excluded from groups and tested individually using either existing criteria for screening or newly developed survey instruments [16, 17].

While CDC guidelines [4] caution against using negative test results “as the sole basis for treatment or other patient management decisions,” testing criteria currently select for symptomatic individuals. Among individuals presently excluded by testing guidelines, we argue that grouped testing is an effective strategy for finding asymptomatic cases that would otherwise be missed.

Overall, we propose a very simple pooling strategy (Figure 1) and thoroughly evaluate the relevance of this test to the COVID-19 pandemic. First, on a 96-well plate, pool each row (8) and each column (12), and perform a total of 20 tests on the pooled samples across the 96 individuals being tested (plus appropriate positive and negative controls). Then, all samples for which both the row and column were positive are re-run in a validation test to determine which individuals are positive. We show below that this strategy works to improve upon existing testing and evaluate the benefits across both simulated and real testing data.

*Stanford University, Stanford, CA; nasa@stanford.edu

†Oracle Corporation, New York, NY; daniel.l.klein@oracle.com

‡Google Inc., Cambridge, MA

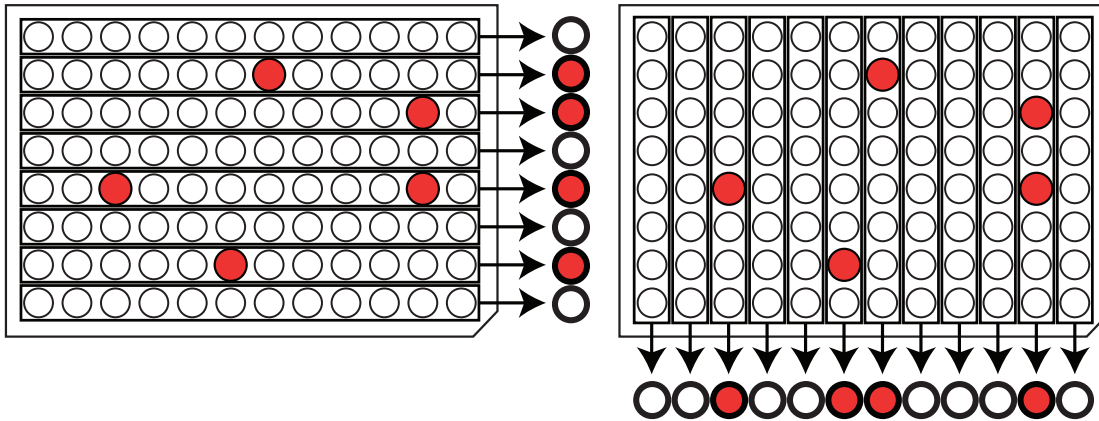
Group testing of 96 samples with 36 tests

Each well contains a single patient sample

● : samples from positive cases

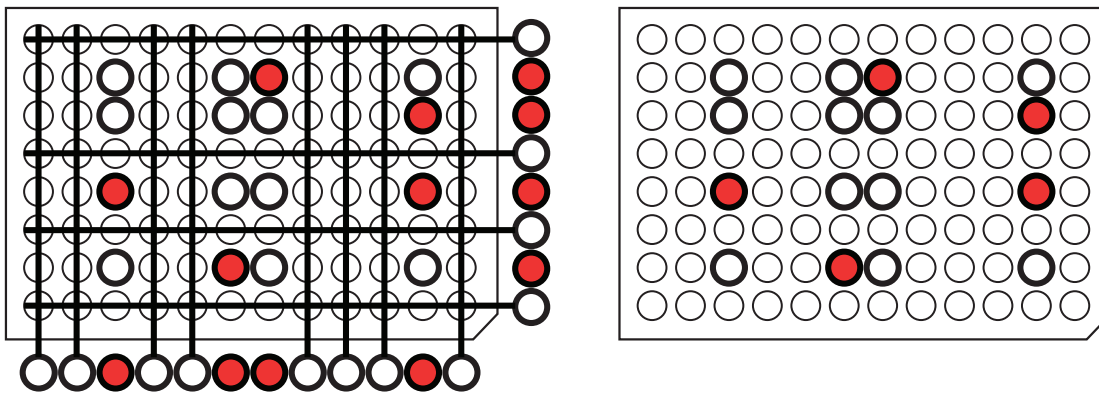
1. Test mixtures of all the samples in each row (8 tests)...

...and mixtures of all the samples in each column (12 tests)



2. Exclude negative rows and columns

3. Test each remaining sample (16 tests)



8 rows + 12 columns + 16 remaining =

36 tests needed in total (2.6x improvement)

Figure 1: Overview of the testing procedure. On a 96-well plate, pools of samples from each row, and separate pools of samples from each column, are tested in a first round. Positive rows and columns are present if any individuals therein are positive, so negative rows and columns can be excluded. The individuals from positive rows and columns are then tested individually to assess their COVID-19 status. Individuals for which both their row and column tested negative are not screened further.

An interactive tool is available at https://technopolymath.shinyapps.io/pooled_covid_testing/

Results

First, we evaluated whether group testing might improve upon existing individual-test methodologies. Testing three pooling strategies — four-well pooling (pooling four adjacent wells in a single dimension), 96-well (8x12) plate column and row pooling, and 384-well (16x24) plate column and row pooling — we found that all improve significantly upon naive testing, and that each performed better for a particular prevalence, as expected (Figure 2). All three required fewer tests than the naive method at prevalence less than 10%. We have implemented these simulations in an R package that is available at <https://gitlab.com/technopolymath/pooled-diagnostics>, and an interactive web visualization is available at https://technopolymath.shinyapps.io/pooled_covid_testing/.

We additionally wanted to estimate the total saved tests relative to those conducted so far, under the hypothetical that pooling had been used throughout the pandemic. Using testing and positive test estimates from Our World in Data [15], we chose an optimal testing strategy per country from between 4-well pooling; 96-well plates; and 384-well plates. We chose the strategy per country which minimized the tests needed at the observed prevalence. This resulted in an expected total of 670,849 tests globally to re-test everyone who has previously been screened for SARS-CoV-2 (1.85M individuals), corresponding to a three-fold improvement in cases per test. Ignoring country-specific prevalence data, 96-well plates and between 5-way and 8-way well pooling resulted in an expected $\approx 780,000$ tests (Table S1). When considering country-specific prevalence, however, the single strategy which performed best on average across countries was the 96-well plate (at just under a third as many tests as standard approaches; Figure S3). These findings suggests that the 96-well plate row and column pooling approach is a straightforward compromise if no information on population prevalence is available.

This approach is even better suited to mild and asymptomatic individuals, as population prevalence is most likely substantially lower than observed in the cases tested by most countries, an intuition borne out in recent modeling efforts [14]. We assume that the prevalence estimates from existing testing are biased due to primarily testing symptomatic individuals, and therefore evaluated our approach at a range of prevalence biases (Figure S2). At up to ten-fold lower prevalence than current tests, our approach still discovers more cases per test than existing diagnosis protocols.

Together, these results suggest that testing mildly symptomatic, or asymptomatic, individuals using a group testing approach will discover the same number of cases using the same number or fewer tests, albeit in a population with much lower prevalence. This enables large numbers of lower-risk individuals to be tested, either from mild symptomatic individuals or contact tracing, which in turn empowers behavioral change.

Discussion

We propose a pooling extension to existing SARS-CoV-2 testing protocols to boost efficiency and better utilize scarce resources. These are not meant to be a replacement for wide-spread testing, but instead might act as a stop-gap measure to reduce the number of RNA extractions and testing steps performed while still identifying asymptomatic or mild cases. Even so, there are numerous reasons why the use of such strategies is not popular.

1. Grouped testing, particularly with an iterative test as here, uses more consumables, such as pipette tips, than traditional methods per test. This is at the benefit of more individuals being tested.
2. More complex sample assignment strategies (e.g. probability driven) can reduce the total number of tests but these are difficult to implement in practice. As such, we frame our proposal in terms of the standard 96-well plates and show that a single plate's worth of pooling already is very effective at expected population prevalence, supporting the use of this strategy (Figure S4).
3. Similarly, the random assignment of individuals to wells, which is much easier than clinical-risk-driven ordering, has nearly the same improvement (Figure S1). In addition, the total improvement in tests between group testing everyone and only group testing low-risk individuals is nearly indistinguishable at testing prevalence below 10%, suggesting that either is an equally useful option to pursue.

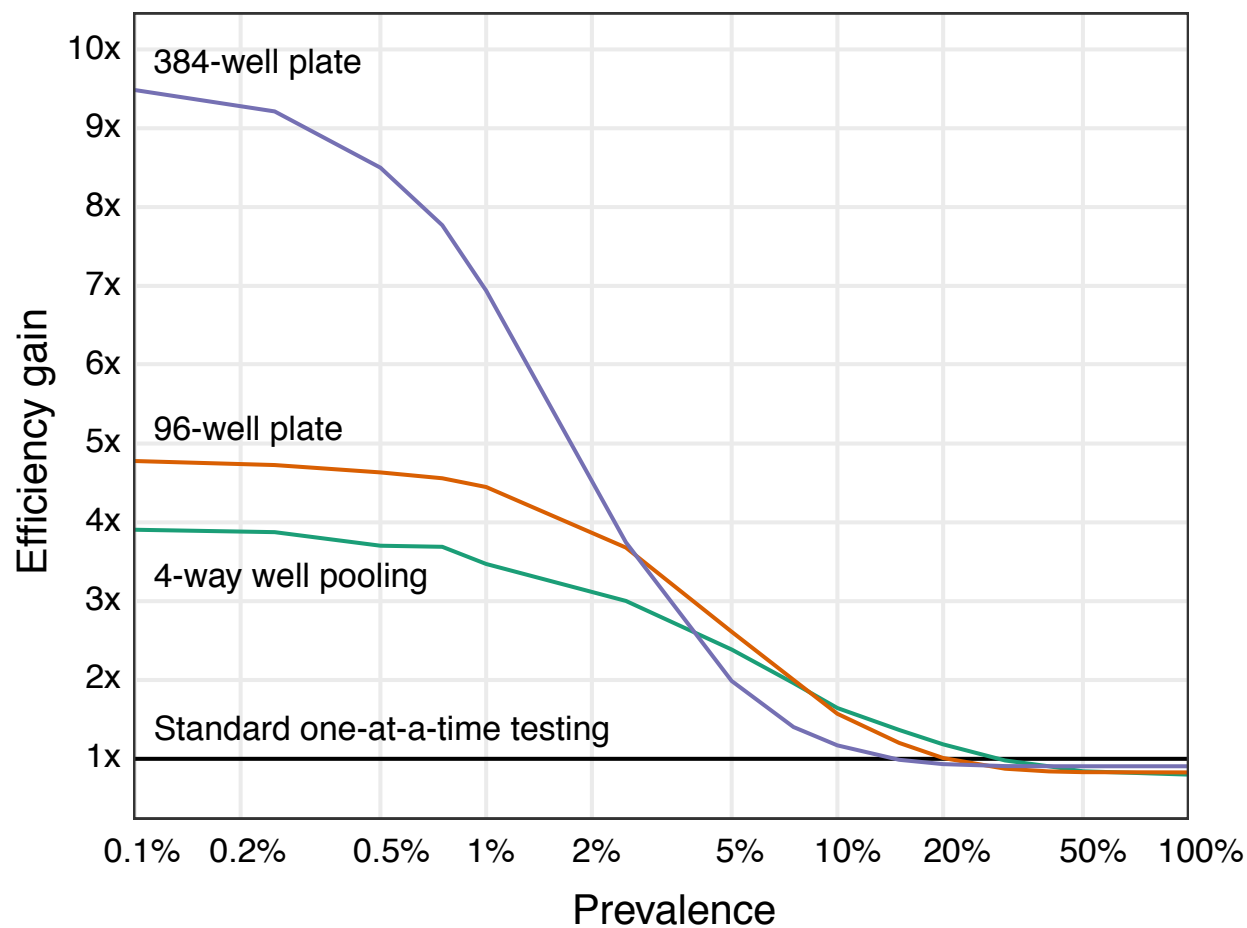


Figure 2: Evaluation of mean improvement of the pooling strategies relative to naive testing under different predicted prevalence. At 2% prevalence and below, row and column pooling on 384-well plates performs best; between 2% and 10%, row and column pooling on 96-well plates performs best; and at greater than 10% prevalence, four-way pooling of wells performs best.

4. Mixing samples from multiple individuals in a single tube risks contamination. We reiterate that this method should not be used as the sole diagnostic criterion for critical ill patients, but rather can be considered to develop epidemiological parameters and/or perform initial screening in low-risk individuals then follow up with confirmatory tests.
5. Interpretation and reporting requirements assume individual testing. In particular, CDC guidelines [4] requires a human RNA positive result (RP+) for an interpretation of “2019-nCoV not detected”. In the pooled strategy, validation tests will be run on a number of negative cases, which may be sufficient to demonstrate a high RP+ rate on this population. Pooled negatives might be considered as “Invalid Results”, and thus should not be reported, until such time as RP+ is confirmed individually.
6. Picking individual samples to re-test is prone to pipetting errors. We developed a strategy with this in mind, as it only requires re-testing samples from any individual row or column which tests positive as a pool. In this way, the workflow disturbance is minimized. See “Minimizing pipetting errors” below.
7. If RNA extraction is the limiting step, this approach requires samples to stay stable for the first qPCR run, in order to allow RNA extraction from the positive samples. We note there is some evidence that RNA extraction is not necessary for qPCR of SARS-CoV-2 RNA [3], though more testing is needed.
8. This method does not remove any clinical or logistical barriers to obtaining large numbers of individual swabs for testing. As such, it relies on the presence of other infrastructure to obtain prevalence estimates. Alternative strategies for prevalence-only modeling, such as wastewater-based epidemiology [7], might be more suitable in settings where individual swabs are not available.
9. At non-negligible prevalence, a substantial number of tests are still needed to cover large populations. The four-fold improvement we observe at prevalence 2% still results in large numbers of total tests, which might remain above capacity.
10. Requiring two stages of qPCR might delay return of results. However, testing resource limitations are also causing delays in the current pandemic, and confirmatory tests are often employed for diagnostics. Finally, for the targeted subset of people who have COVID-19 but do not have severe symptoms, delays may have limited clinical significance or consequences for public health interventions.

We encourage further work into better approximate methods for obtaining prevalence estimates through more sophisticated pooling strategies, including using the virus concentration from qPCR (in either equally or unequally pooled samples). In addition, we see great potential of pooling for repeat testing of the same population of individuals on sequential days or weeks. Hospitals, government offices, corporations, and educational institutions, across which many individuals cannot work from home [2], require self-isolation following positive results. This decreases the case counts below the population average on subsequent days and increases the efficiency of group testing strategies. Among individuals with low risk of serious complications, this methodology could use non-qPCR based assays. For instance, RT-LAMP assays [9, 18, 21, 10] might be particularly well-suited to rapid, first-line screening methodologies at scale. This strategy might also be applied to seropositivity screening once reliable tests are available for COVID-19, though rigorous testing of the pooling of serum is necessary to evaluate this more thoroughly.

Further extensions of this approach might be justifiable given the current extreme limitations on testing resources. For instance, testing efficiency can be increased at the expense of specificity by pooling people in close contact (ex. family groups, work units) into a single sample prior to group allocation or RNA extraction. This reduces test volume by a constant factor under the assumption that individuals in close contact have positively correlated infection status and that cluster identification is more vital to suppression than determining a particular individual’s infection status. In addition, pooling samples from the same individual (sputum, nasopharyngeal swab, etc) will obtain similar benefits with essentially no loss of information.

Overall, we present a straightforward and directly applicable approach to extend testing for SARS-CoV-2 to lower-risk individuals using pooling. While there are numerous technical hurdles left to overcome, we anticipate that strategies such as those proposed here will be valuable in obtaining more accurate estimates of population prevalence and disease spread in the coming months.

Minimizing pipetting errors

Pipetting errors are clearly a large concern in the implementation of this approach, as single missteps are both difficult to detect and result in both false positives and false negatives. We consider multiple possible mitigation strategies:

1. Multi-channel pipettes enable simultaneous multi-sample manipulation by hand. Given the gridded nature of our testing system, multi-channel pipettes are likely effective. Multi-channel pipetting into a deepwell plate for pooled sample storage is likely the simplest option. For example, into a 2.5 mL deepwell plate, pipette 150 μ L of each pre-extraction row and column sample to yield a total of 1.8 mL of pool for rows and 1.2mL of pool for columns. Then mix thoroughly and extract from standard volumes. We note that this mixing increases risk of operator exposure, so strict biosafety precautions must be adhered to.
2. Many SARS-CoV-2 testing laboratories are using pipetting robots to rapidly and accurately scale RNA extractions and assays to large numbers of individuals. While protocols do not exist to our knowledge for group testing on these platforms, they are ideally suited to tasks such as this. We propose such a strategy should be validated using traditional pipetting before using valuable robotic platform resources.
3. Finally, in the absence of both robotic and multi-channel automation, pipette guiding tools such as iPipet [22] can be used for suggesting pipetting steps in a fixed order. This reduces the possibility of pipetting errors without requiring a multi-channel pipette, but does require tablet access.

Independent of the modest efficiency gains of clinical-risk-driven ordering (as in Figure S1), ordering of individual samples so that pipetting order is from low- to high-risk samples minimizes the impact of unintended transfers. In the example strategy in Figure 1, this would be achieved by placing high-risk samples in the bottom-right.

Competing Interest Statement

The authors declare no competing interests.

Acknowledgments

We would like to thank Xinchun Wang, Hanna Ollila, Diego Calderon, Courtney Smith, Josh Tycko, Veronica Casas, Michael Wainberg, Jonathan Pritchard, Ruth Brillman and Christopher Davis for giving comments on this work. In particular, Courtney Smith suggested the use of pooling to evaluate individuals post infection, Michael Wainberg provided advice on figure design, and Xinchun Wang and Diego Calderon altered us to recent references. N.S.-A. is supported by a Stanford Graduate Fellowship. We thank Our World in Data for providing and updating the global testing effort data. We thank numerous government agencies for providing up-to-date information on COVID-19.

References

- [1] Andrea Crisanti and Antonio Cassone. *In one Italian town, we showed mass testing could eradicate the coronavirus*. <https://www.theguardian.com/commentisfree/2020/mar/20/eradicated-coronavirus-mass-testing-covid-19-italy-vo> [Accessed: 2020-03-24]. March 20, 2020.
- [2] Marissa G Baker. “Characterizing occupations that cannot work from home: a means to identify susceptible worker groups during the COVID-19 pandemic”. Mar. 2020.
- [3] Emily A Bruce et al. “RT-qPCR Detection of SARS-CoV-2 RNA from patient nasopharyngeal swab using QIAGEN RNEasy kits or directly via omission of an RNA extraction step”. In: *bioRxiv* (2020).

- [4] Centers for Disease Control and Prevention, Division of Viral Diseases. *CDC 2019–Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel*. <https://www.fda.gov/media/134922/download> [Accessed: 2020-03-23]. March 15, 2020.
- [5] Robert Dorfman. “The detection of defective members of large populations”. In: *The Annals of Mathematical Statistics* 14.4 (1943), pp. 436–440.
- [6] Yaniv Erlich et al. “DNA Sudoku - Harnessing high throughput sequencing for multiplexed specimen analysis”. In: *Genome Research* 19.7 (2009), pp. 1243–1253.
- [7] Iria González-Mariño et al. “Spatio-temporal assessment of illicit drug use at large scale: evidence from 7 years of international wastewater monitoring”. In: *Addiction* 115.1 (Jan. 2020), pp. 109–120.
- [8] Paul Hammel. “.@GovRicketts says state is ‘pooling’ tests now, so combining five tests into one test tube. If it comes back ‘negative,’ then you know all were negative; if one was ‘positive,’ then those five will have to be retested. So that expands capacity to 400+ tests a day by public labs. <https://twitter.com/PaulHammel0WH/status/1242539326568902661> [Accessed: 2020-03-24]. March 24, 2020.
- [9] Minghua Jiang et al. “Development and validation of a rapid single-step reverse transcriptase loop-mediated isothermal amplification (RT-LAMP) system potentially to be used for reliable and high-throughput screening of COVID-19”. In: *medRxiv* (2020).
- [10] Laura E Lamb et al. “Rapid Detection of Novel Coronavirus (COVID-19) by Reverse Transcription-Loop-Mediated Isothermal Amplification”. In: *medRxiv* (2020).
- [11] Andrea Navarro. “Mexico’s Low Coronavirus Count Spurs Doubts On Testing Rate”. In: *Bloomberg News* (Mar. 2020).
- [12] Maria Orlova. *Government of Kyrgyzstan announces number of coronavirus test systems*. https://24.kg/english/146493_Government_of_Kyrgyzstan_announces_number_of_coronavirus_test_systems/. Accessed: 2020-3-26. Mar. 2020.
- [13] RM Phatarfod and Aidan Sudbury. “The use of a square array scheme in blood testing”. In: *Statistics in Medicine* 13.22 (1994), pp. 2337–2343.
- [14] Lionel Roques et al. “Mechanistic-statistical SIR modelling for early estimation of the actual number of cases and mortality rate from COVID-19”. Mar. 2020.
- [15] Max Roser, Hannah Ritchie, and Esteban Ortiz-Ospina. “Coronavirus Disease (COVID-19)—Research and Statistics”. In: *Our World in Data* (2020).
- [16] Hagai Rossman et al. “A framework for identifying regional outbreak and spread of COVID-19 from one-minute population-wide surveys”. In: *medRxiv* (2020).
- [17] Arni S Srinivasa and Jose A Vazquez. “Identification of COVID-19 Can be Quicker through Artificial Intelligence framework using a Mobile Phone-Based Survey in the Populations when Cities/Towns Are Under Quarantine”. In: *Infect. Control Hosp. Epidemiol.* (2020), pp. 1–18.
- [18] Mohamed El-Tholoth, Haim H Bau, and Jinzhao Song. “A Single and Two-Stage, Closed-Tube, Molecular Test for the 2019 Novel Coronavirus (COVID-19) at Home, Clinic, and Points of Entry”. In: (2020).
- [19] Idan Yelin. *Pool testing for COVID-19 works! Can detect a single positive sample within a pool of 64 samples. Collaboration of: @RambamHCC @TechnionLive @BiologyTechnion @RoyKishony @yuval_geffen*. <https://twitter.com/idanyelin/status/1239930847748227079> [Accessed: 2020-03-23]. March 17, 2020.
- [20] Idan Yelin et al. “Evaluation of COVID-19 RT-qPCR test in multi-sample pools”. In: *medRxiv* (2020). DOI: 10.1101/2020.03.26.20039438. eprint: <https://www.medrxiv.org/content/early/2020/03/27/2020.03.26.20039438.full.pdf>. URL: <https://www.medrxiv.org/content/early/2020/03/27/2020.03.26.20039438>.

- [21] Lin Yu et al. “Rapid colorimetric detection of COVID-19 coronavirus using a reverse transcriptional loop-mediated isothermal amplification (RT-LAMP) diagnostic platform: iLACO”. In: *medRxiv* (2020).
- [22] Dina Zielinski et al. “iPipet: sample handling using a tablet”. In: *Nature methods* 11.8 (2014), p. 784.

Supplement

We provide an overview of the simulation methods employed. We evaluated models using a sampling approach, drawing samples from either a Bernoulli distribution with a probability equal to the assumed prevalence, or a beta-Bernoulli distribution with α parameter 5 and β parameter $5/prevalence$ to account for inter-individual variation in risk. All draws were independent. These were averaged over 1000 replicates to smooth over any sampling noise; sample sizes are adjustable in our web interface.

Next, the draws are used to assign true case status for each simulation run. The number of random values thus corresponds directly to the count of cases tested by a given method in a single batch. Of note, the draws from the beta are used as the case probability directly for the row major ordered case prevalence tests (Figure S1), which are only informative under the beta-Bernoulli simulations.

Finally, the results with our two-stage approach (pooled screening followed by individual-test followup of potential cases, and optional confirmation of positive status) are compared with corresponding naive screening (individual-test screening followed by optional and default individual-test confirmation of cases). Evaluating different levels of test stringency for both the grouped and naive strategies is part of the web interface.

We assume for all methods that samples are distributed at random and independently. However, sample collection for this approach might be conducted through contact tracing and door-to-door sampling, where we expect co-habitation. In these cases, the assumption of lack of correlation of test status between individuals does not hold. We anticipate relevant extensions to this work might address this concern.

While the random assignment might at first seem inefficient, we evaluated a number of additional strategies, including ones which use the case probability information, and found that they did not perform meaningfully better than random assignment (Figure S1). As such, given the ease of the random assignment process, we recommend that as a default. Comparison to case ordered alternatives is available on the web interface.

Strategies that pool across multiple plates offer further test efficiency gains, depending on the prevalence. Because this dependence is sensitive to prevalence and because using a single 96-well plate is near-optimal at an expected 5% prevalence (Figure S4), we recommend this practically simpler strategy.

Sample stratification provides another opportunity for realizing efficiency gains. By separating samples from symptomatic individuals from others, we can artificially decrease prevalence in the individuals tested using pooling, which decreases the tests per individual. This makes it easier to find asymptomatic carriers by expanding test coverage, though practical trade-offs between case abundance and overall testing capacity are worthy of consideration.

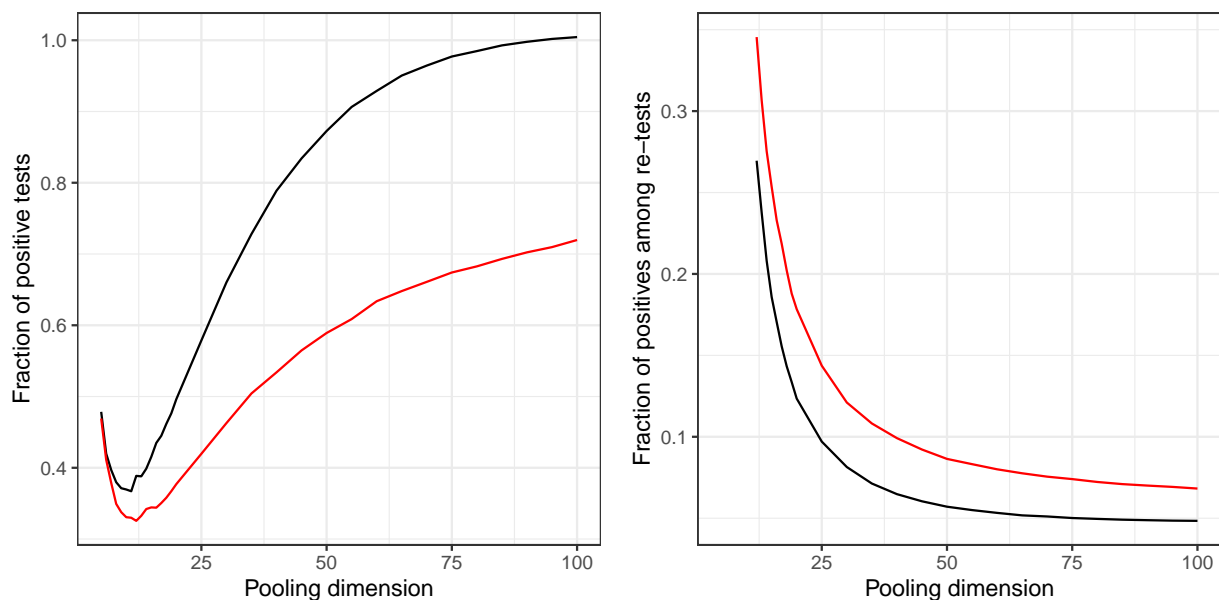


Figure S1: Comparison of random (black) versus case-probability-ordered (red) modes of pooling. Simulations done with known case probability drawn from $Beta(0.5, 10)$ (5% prevalence). Pooling dimension k is the size of the $k \times k$ plate of individuals to be tested (rows and columns). In both the total number of tests relative to naïve (left) and the fraction of positives in the resulting secondary testing scheme (right), the ordered method results in slight improvements around the optimum pooling dimension. We suggest these are not sufficient to warrant the additional complexity this method entails.

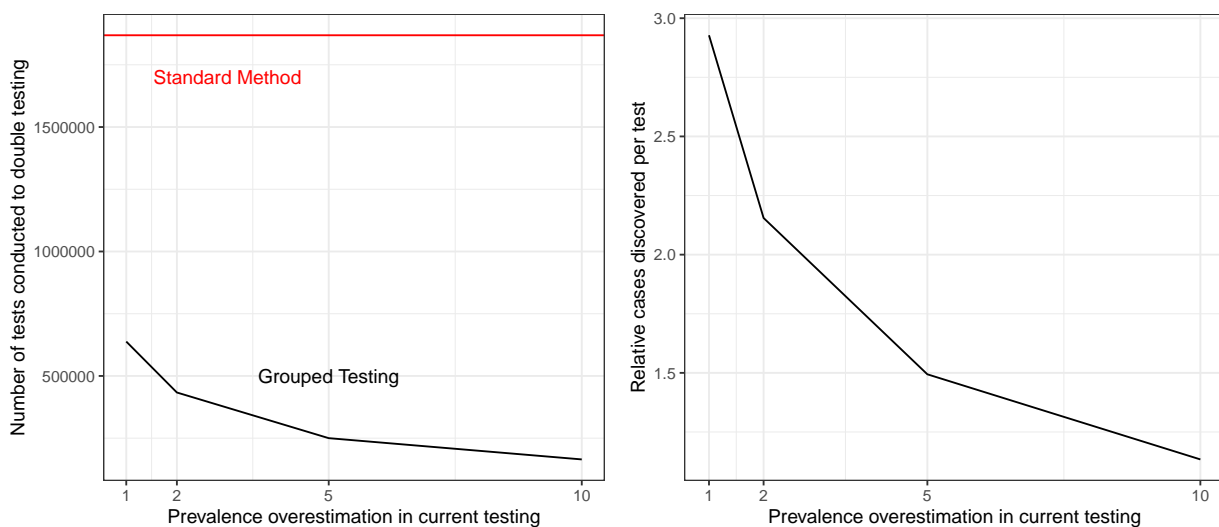


Figure S2: Expected improvements relative to current testing. Across global tests performed so far, we choose the best grouped testing strategy per country. In order to double the current tests performed (left), grouped testing requires progressively fewer total tests at lower prevalence. Prevalence scaled by observed prevalence within each country (so 2-fold overestimation in a country with 3% observed prevalence = 1.5% tested prevalence). Testing these lower-prevalence individuals using group testing finds more cases per test than existing strategies did, even at lower overall prevalence than existing screening (right).

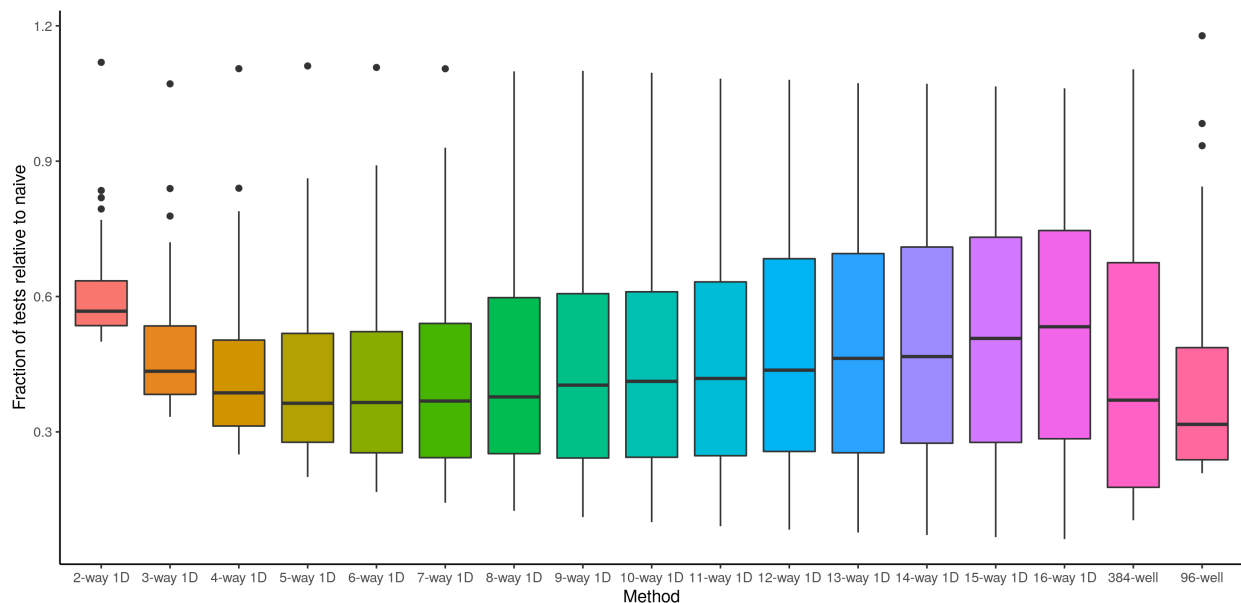


Figure S3: Comparison of pooling strategies across countries used in the optimization simulation. Across all 60 sampled countries [15], the 96-well plate row and column approach had the largest median improvement (just greater than three-fold). Alternative strategies that work well across a large number of countries are 384-well plates and 5-way to 8-way well pooling.

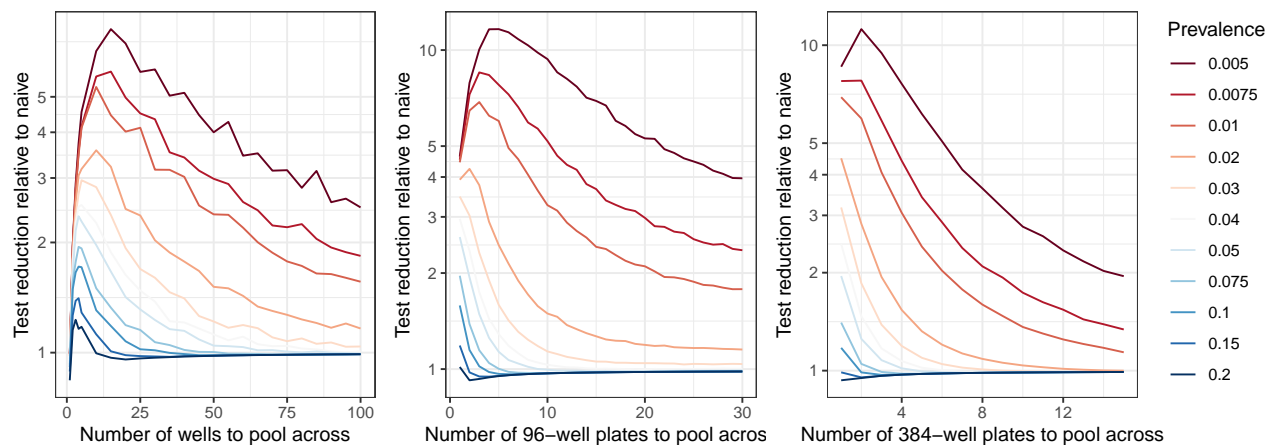


Figure S4: Comparison of prevalence scaling of efficiency for simple well pooling (left), 96-well plates (middle), and 384-well plates. Simple well pooling is more efficient at prevalence $\geq 10\%$, while 384-well plates are more efficient for prevalence $\leq 1\%$. Pooling of 3-5 wells and single 384-well plates is recommended at prevalence outside the range that is well captured by a single 96-well plate.

Country	COVID-19 cases	COVID-19 tests	COVID-19 prevalence	# tests with standard	# tests with 4-way pool	# tests with 96-well pool	# tests with 384-well pool
Armenia	78	813	0.096	891	465	494	680
Australia	709	113,615	0.006	114,324	31,130	24,710	13,830
Austria	2,196	15,613	0.141	17,809	10,913	12,553	15,514
Bahrain	269	18,645	0.014	18,914	5,891	4,391	3,249
Belarus	27	16,000	0.002	16,027	4,128	3,361	1,711
Belgium	1,243	18,360	0.068	19,603	8,959	8,662	12,055
Brazil	77	2,927	0.026	3,004	1,047	798	804
Canada	846	113,121	0.007	113,967	31,334	24,930	14,526
Colombia	128	4,103	0.031	4,231	1,456	1,227	1,327
Costa Rica	69	1,039	0.066	1,108	495	477	67
Croatia	81	1,264	0.064	1,345	632	575	793
Czech Republic	694	11,619	0.060	12,313	5,530	4,992	6,956
Denmark	1,151	10,730	0.107	11,881	6,716	7,164	9,513
Estonia	267	2,504	0.107	2,771	1,552	1,659	2,223
Faroe Islands	72	1,641	0.044	1,713	681	576	748
Finland	369	3,000	0.123	3,369	2,019	2,263	2,822
France	4,499	36,747	0.122	41,246	24,289	26,966	34,831
Germany	3,795	167,000	0.023	170,795	60,120	43,473	40,504
Hungary	85	3,007	0.028	3,007	1,025	860	892
Iceland	330	9,189	0.036	9,519	3,611	2,872	3,446
India	191	14,514	0.013	14,705	4,267	3,383	2,374
Indonesia	227	2,028	0.112	2,255	1,310	1,391	1,838
Iran	11,364	80,000	0.142	91,364	57,200	63,828	79,782
Ireland	223	6,600	0.034	6,823	2,508	2,044	2,343
Israel	427	10,864	0.039	11,291	4,475	3,648	4,421
Italy	41,035	206,886	0.198	247,921	173,784	203,452	221,376
Japan	873	14,901	0.059	15,774	6,809	6,174	8,631
Latvia	86	3,205	0.027	3,291	1,124	897	902
Lithuania	48	1,154	0.042	1,202	469	399	499
Malaysia	900	13,876	0.065	14,776	6,480	6,418	8,765
Malta	9	889	0.010	898	264	199	127
Netherlands	128	6,000	0.021	6,128	2,070	1,566	1,388
Norway	1,552	43,735	0.035	43,287	16,225	14,025	15,990
Pakistan	302	1,979	0.153	2,281	1,450	1,669	2,018
Palestine	38	2,519	0.015	2,557	765	602	448
Panama	86	1,455	0.059	1,541	678	632	851
Philippines	230	1,269	0.181	1,472	1,001	1,185	1,339
Poland	355	13,072	0.027	13,427	4,640	3,669	3,698
Qatar	439	8,400	0.052	8,839	3,847	3,339	4,424
Romania	277	8,284	0.033	8,561	3,098	2,540	2,881
Russia	147	143,519	0.001	143,666	36,884	30,062	15,146
Slovakia	123	2,707	0.045	2,830	1,155	975	1,253
Slovenia	286	9,860	0.029	10,146	3,648	2,862	2,985
South Africa	150	6,438	0.023	6,588	2,118	1,720	1,611
South Korea	8,652	316,664	0.027	325,316	113,999	89,939	89,745
Spain	11,178	30,000	0.373	30,000	41,178	35,337	33,107
Sweden	1,121	14,300	0.078	15,421	7,350	7,598	10,441
Switzerland	209	4,000	0.052	4,209	1,760	1,576	2,116
Taiwan	108	21,376	0.005	21,484	5,664	4,600	2,497
Thailand	177	7,084	0.025	7,261	2,408	1,919	1,888
Ukraine	26	316	0.082	342	170	173	240
United Arab Emirates	86	125,000	0.001	125,086	31,625	26,115	13,128
United Kingdom	2,630	64,621	0.041	67,251	26,171	21,796	26,756
United States	9,415	103,945	0.091	113,360	56,546	60,729	83,696
Vietnam	85	15,637	0.005	15,722	4,315	3,387	1,859
Sum of above	110,168	1,858,034	0.061	1,858,034	821,420	782,851	817,661
Hong Kong*	4	5,271	0.001	5,275	1,344	1,102	553
Kyrgyzstan*	0	1,545	0	1,545	386	321	160
Mexico*	12	278	0.043	290	113	96	123
New Zealand*	11	584	0.019	595	189	146	122
Turkey*	0	2,900	0	2,900	725	604	302

Table S1: Comparison of pooling strategies across countries and choice of the 96-well plate as a recommended strategy. Across 55 countries with available matched case and testing data [15], the estimated number of tests required to duplicate the existing testing effort (at country-observed prevalence among tested individuals) exclusively using the given method is provided. When examining how many countries are optimized by each strategy, however, 40% (24 of 55) of countries save the most tests using the 96-well plate strategy. Choosing the best strategy per country results in a total of 670,849 tests, the vast majority of which are saved by the 384-well plate and 96-well plate strategies. Data from different countries are across a range of dates (3/5-3/20) as number of tests are not systematically reported.

* Case counts were not recorded by [15] on the same day as test counts for Turkey, Kyrgyzstan [12], Hong Kong, Mexico [11], and New Zealand, are obtained from source and [15], and are excluded from global estimates. We infer zero cases in Turkey from the date reporting tests (3/10/2020) and first case date (3/12/2020) and exclude two probable cases in New Zealand.