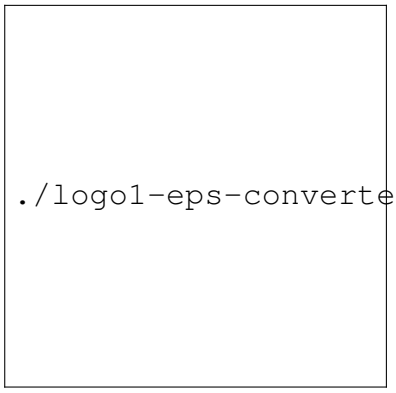


Ordered Rectangular Pooling for Improved Public Health Testing

Terrance E. Boult [1,*], **Yanyan Zhuang** [1,*], **Keith Paulson** [1,2†], **and Crystal Vander Zanden** [3]

[1] *Department of Computer Science, University of Colorado Colorado Springs, Colorado Springs, CO, USA*
[2] *MITRE Corporation, Colorado Springs, CO, USA*
[3] *Department of Chemistry & Biochemistry, University of Colorado Colorado Springs, Colorado Springs, CO, USA*

Correspondence*:
Terrance E. Boult
Department of Computer Science, University of Colorado Colorado Springs,
Colorado Springs, CO 80918, USA
tboult@vast.uccs.edu

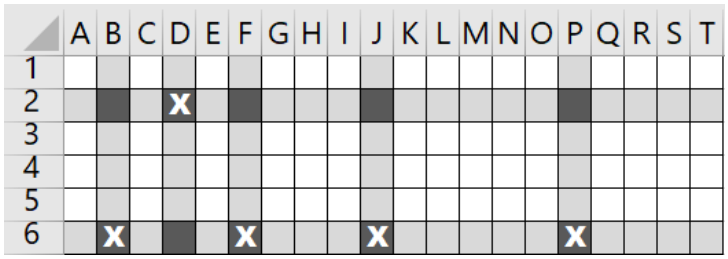
Yanyan Zhuang
Department of Computer Science, University of Colorado Colorado Springs,
Colorado Springs, CO 80918, USA
yzhuang@uccs.edu

ABSTRACT

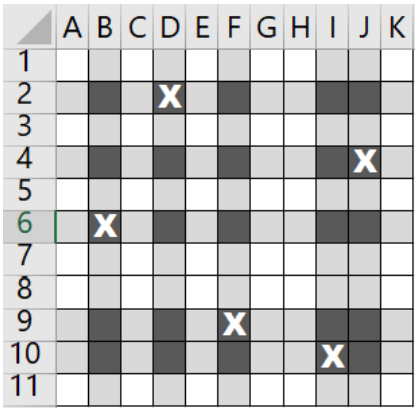
COVID19 has challenged the current day's medical testing capability. Pooled testing is an effective approach that can reduce the overall costs. Current pooled testing techniques usually make the assumption that the entire population has a uniform probability of infection. However, symptomatic and asymptomatic patients have vastly different probabilities of testing positive. In this paper, we introduce a novel pooled testing technique that can handle non-uniform likelihood of infection, which produces pool sizes that suit different sub-populations. Our simulation results show that our approach improves the performance over standard double-pooled designs by 2%-16%, by 5%-125% over single-pooled testing, and by 25%-500% over classic individual testing.

Keywords: Pooled testing, Dorfman, COVID19 testing, population testing, non-uniform likelihood of infection

[†] Keith Paulson's affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions, or viewpoints expressed by the author. Approved for Public Release; Distribution Unlimited. Public Release Case Number 20-2147. Applied to release/transfer Copyright 2020 The MITRE Corporation.



(1a) Ordered Rectangular Pool.



(1b) Standard Double Pool.

Figure 1. Comparison of the ordered rectangular-pooled testing and standard double-pooled testing. The core idea comes from the fact that we have two conflicting objectives – rejecting negatives as much as possible, but also reducing ambiguity so we reduce the number of individual tests needed. In both figures, an X indicates an underlying positive sample that is to be detected via testing, and light gray indicates rows/columns that test positive. When both a row and a column are positive, the individual cell at the intersection (darker gray) needs to be tested in the second round. Consider a population of 120 subjects with an average positivity of 4.2%, of which there are 100 asymptomatic cases with 1% expected positivity, and 20 symptomatic/higher-risk cases with an expected positivity rate of 20%. For ordered rectangular pooling, this results in a 6-row by 20-column design (1a). After sorting the samples according to approximate positivity, we fill rows 1-5 in (1a) with samples from the low-positivity asymptomatic group (one positive sample expected), and fill the last row with the samples from higher-positivity symptomatic subjects (four positive samples expected). The ordered rectangular testing in (1a) requires only 36 tests in total – 20 column-pooled tests, 6 row-pooled tests, and 10 individual tests for each of the darker cells. The standard double-pooled testing for the same population would use an 11x11 matrix (1b). It uses only 22 pooled tests – 11 column-pooled tests and 11 row-pooled tests. However, because the samples are un-ordered, it is expected to need 25 individual tests (dark gray cells in (1b)), for a total of 47 tests. While not shown, using the optimal Dorfman-style testing would use a pool size of 4 and would require 49 tests.

1 INTRODUCTION

COVID19 has increased the need for rapid testing. In early testing within the US, the focus was on symptomatic at-risk patients. However, as the number of cases increases, there is a growing need to test all patients, especially asymptomatic patients. A few organizations have been exploring the Dorfman-style (1) single-pooled SARS-Cov-2 (COVID19) testing (2, 3, 4), where material from multiple subjects are pooled together. The combined samples are tested, and if negative, all samples are considered negative; if the combined pool is positive, then all individuals are tested separately. Studies have shown that pooled testing can be used for COVID19 testing, and the US FDA is now allowing its usage under emergency-use authorization (**?**). Recent research also shows that pooling samples post RNA extraction with a pool size up to 64 is possible (3). Further, using raw samples, pools of sizes up to 10 (the max size tested by the authors of (4)) were also effective with no loss in sensitivity. The use of raw sample pooling can save significantly on the supplies and time needed for RNA extraction. Interestingly, there are also results showing that one can obtain up to five $200\mu L$ samples from a single nasopharyngeal swab (5), allowing three to five pooled tests from just a single sample.

Previous research has developed multi-pooled testing, also called matrix testing or group testing, like the examples shown in Figure 1. A sample is put into multiple pools, i.e., a column-pool and a row-pool in a matrix. When a sample is tested positive in both a row- and column-pool, it is considered a potential

positive case and will be tested individually in the second round. During pooled testing, if there is more than one positive row or column, then all the intersections require individual testing to resolve which ones are positive. For example, in Figure 1b, 5×5 individual tests are needed (shown as dark gray cells) to detect the five underlying positive cases (indicated by an X in the figure). The majority of such multi-pooled testing techniques have used pools of the same size (6, 7, 8, 9, 10). Hence, we call these square-pool testing, with the pools as rows and columns in a square matrix. While the authors of (11) considered unequal (rectangular) matrix pools, they concluded that "matrix methods are most efficient when the matrices are square."

Pooled sampling can be used for statistical analysis of populations, as well as for efficient detection of individuals that match the testing criterion. As mentioned above, prior work on pooled testing for virus detection presumes a single probability of infection applied to everyone in the population. However, assuming a single-population probability is unrealistic. For example, someone who has confirmed contact with an infected person or someone that is symptomatic has a higher chance of testing positive. Asymptomatic patients without contacts with known positive cases have lower risks of infection. COVID19 testing sites have inherently mixed populations, and many testing sites have a much higher overall positivity rate than 10%, where the standard double-pooled testing shows limited benefits.

Pooled testing for detection seeks to minimize total test costs with two different objectives. It involves rejecting negative samples quickly while reducing the need for individualized disambiguation testing needed when multiple rows/columns test positive. We address these two objectives in a novel way, using a rectangular-double-pooled approach where we sort items to ensure as many rows as possible have low-probability of positivity to reject negatives, while the columns have a near-uniform probability of positivity to reduce the need for individual exclusion testing, as shown in Figure 1a. In this paper, we show that **if one assumes that subsets of data have a non-uniform likelihood of infection, then, not surprisingly, different pool sizes should be designed for subsets of data, and ordered rectangular pools provide an easy way to reduce costs.**

Our novel approach to double-pooled designs for non-uniform probabilities of infection significantly improves test efficiency and is straightforward to implement. In the range of positivity normally seen in COVID19 testings, our approach improves the performance over standard double-pooled designs by 2%-16%, by 5%-125% over single-pooled testing, and by 25%-500% over classic individual testing (Section 3). The code for computing the optimized designs is publicly available, and code for managing and tracking samples through the pooled testing process is under development.

2 METHOD

The intuition behind our approach is that, assuming the sensitivity of a test is not a strong function of the pool size, then a low probability of infection results in a larger optimal pool size, while a higher probability of infection requires a smaller pool size, potentially even individual testing. By sorting and placing the samples according to their approximate likelihood of infection in row-major order in the matrix, we cluster the low-probability samples into low-numbered long rows, e.g., rows 1-5 in Figure 1a, and higher-probability samples into high-numbered rows, e.g., row 6 in Figure 1a. This also results in a similar overall probability in the columns of the rectangle. Referring back to Figure 1, each column in Figure 1a has a 25% likelihood of being positive, well below the likelihood in a standard double-pooled testing of 11 * 4% = 44% per column (Figure 1b).

We develop a general approximation of the expected cost of testing, based on per-sample probabilities. We note that the data one can obtain during sample collection, such as contact status, symptomatic or not, age, activities, and region/zonal locations, can provide an estimated a priori probability for each sample. Given these positivity estimates, we can estimate the costs for any proposed rectangular-pooled design. In the supplementary material we present some examples as well as the background on deriving the formulas below.

Let the per-sample probabilities be $p_1, \ldots p_N$, where N is the total number of samples. The rectangular-pooled testing fills these in a rectangular matrix with r rows and c columns, where $N \leq r * c$. First, we sort the samples based on the estimated positivity, and then fill the matrix in row-major order. If we assume that there are $\hat{r}$ rows and $\hat{c}$ columns with positive pooled testing responses, then we can approximate the expected total cost $\widehat{TC}$ as a function of r and c. Letting $r + c$ be the cost of all pooled tests, and $E(\hat{r} * \hat{c})$ be the expected cost of all individual tests, we have:

$$\widehat{TC}(r, c) = r + c + E(\hat{r} * \hat{c}) \quad \approx \quad r + c + E(\hat{r}) * E(\hat{c}) \tag{1}$$

where

$$E(\hat{r}) = \sum_{i=1}^{r} \left(1 - \prod_{k=1}^{c} (1 - p_{i*c+k})\right) \quad \text{and} \quad E(\hat{c}) = \sum_{i=1}^{c} \left(1 - \prod_{k=1}^{r} (1 - p_{i+k*r})\right) \tag{2}$$

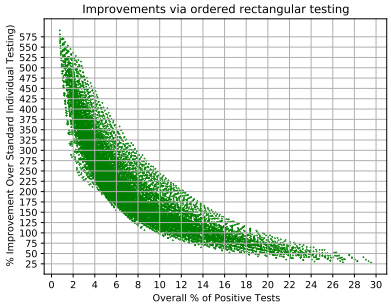
subject to $r * c \geq N$. Given data, this expected value can be quickly optimized by searching over the space where $1 < r, c < 64$. The per-item cost would be $\frac{\widehat{TC}(r,c)}{\#samples}$, where $\#samples$ $(< N)$ is the actual number of samples used to fill the matrix during a test. The details of the derivation are provided in the supplementary material.

Note that this is an approximation, as the expected value of a product $E(\hat{r} * \hat{c})$ may not equal the product of the expected values $E(\hat{r}) * E(\hat{c})$ when the data is not independent. Since we have the same underlying probabilities p_i contributing to both the expected values of positive rows $E(\hat{r})$ and columns $E(\hat{c})$, they are not independent. In simulations, we find that this is a good approximation when the probabilities are low, but becomes increasingly noisy when some probabilities are above 0.3. When the error is high, it generally underestimates the true cost for mixed and higher probabilities. *Importantly, such an approximation error only impacts the estimated cost savings, not the accuracy of the final testing results.*

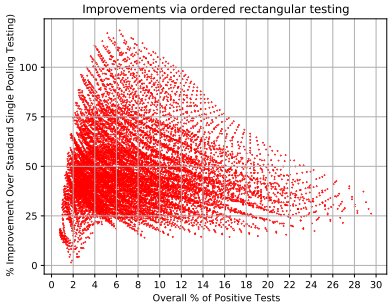
3 RESULTS

In Figure 2, we show the relative percent improvement of the ordered rectangular-double-pooled testing over the standard individualized testing (2c), single-pooled testing (2d), and the optimal square-double-pooled testing (2e). The results show savings over 9,000 different configurations, with the overall population positivity rates between 1% and 30%. Compared to individualized testing, the relative savings are between 25% and 500%. Compared to single-pooled testing, the relative savings are between 2% and 125%. And finally, compared to the optimal square-double-pooled testing, the relative savings are between 2%-16%, with greater savings at lower positivity rates, or with greater variation in the underlying population. Using a two-sided Wilcoxon-non-parametric test, we conclude that these results are significant ($p < .0001$).

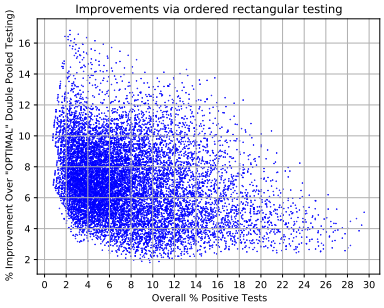
The proposed ordered rectangular-double-pooled testing has two novel aspects: ordering and rectangular pools. Without the ordering, rectangular pools can be less efficient than square pools, as they can degenerate to a single pool. Thus, the use of ordering is critical. What happens to standard square-double-pooled testing if we use ordered data? Not surprisingly, ordering improves the expected performance significantly.



(2c) Improvement over individual testing.



(2d) Improvement over single-pooled testing.



(2e) Improvement over "optimal" square-double-pooled testing.

Figure 2. Performance improvements of ordered rectangular-pooled testing with over 9,000 samples from a two-sub-population model, where one group is asymptomatic with a positivity rate of p_A between .005 and .02, and the other group is symptomatic with a positivity rate of p_S which varies from .1 to .5 Left (2c) is improvement over standard individual testing. Middle (2d) is improvement over optimal single-pooled testing, such as the tests currently allowed by the FDA. Right (2e) is improvement over "optimal" square-double-pooled testing which uses a single overall population positivity rate.

Using a two-sided Wilcoxon-non-parametric test, with Bonferroni, an ordered square-double-pooled testing is significantly better than randomized square-double-pooled testing ($p < .0001$). Additionally, we also find that an ordered rectangle is better than an ordered square ($p < .0001$).

4 DISCUSSION

While we presented an initial model that optimizes ordered rectangular pools for multiple populations, the final model, with per-sample probability estimates, optimizes for any population. The per-sample approach degrades gracefully with increasing error in the probability estimates. The resulting design ensures the correctness of the testing, even though it is non-optimal in terms of the number of tests conducted. In this section, we discuss other issues with pooled testing, some of which are not limited to our approach.

Design assumptions. While we have derived the expected cost allowing optimization of the multi-group ordered rectangular-double-pooled sampling, and experimentally showed it was, in general, better, it is important to understand the assumptions and design accordingly. The first is the assumption of the probability model. We recognize that any probability estimate is unlikely to be accurate, so over-optimization must be balanced. Not only is this the inherent issue of random perturbations in any probability model, but the estimates are also likely to contain bias, since people may intentionally or unknowingly choose not to answer health questions accurately, and often simply do not know about their contacts.

Thus, our actual system estimates the expected costs, and compares it with the expected cost of single pooling and individual testing. While the system can effectively optimize for any mixture of probabilities, via simulation, we find that the ordered rectangular-pooled testing has a large variance and can sometimes cost more than individual testing when probabilities of infection are above 30%. Therefore, our software system will not insert samples into ordered rectangles if their probability is higher than 30%; such samples should be tested individually.

Probability estimates. While ordering is required for the rectangular-pooled testing to be better, of the two sources of relative improvement gain, on average, about 40% is from the sorting of probabilities, and 60% from relaxing the equal pool size constraint. If the probability estimates are terrible, e.g., inverted

labels between groups, the rectangular ordering can be less efficient than the square. Thus, if the probability estimates are questionable, our software can be configured to produce only the ordered square designs. This reduces the maximum efficiency gains, but eliminates any chance of efficiency loss from over-optimization using bad probability estimates.

In a recent paper (12), Friedman et. al. made a compelling argument for "zonal social-distancing" policies that organize people into zones, i.e., a particular structure of groups, and suggested it could be combined with pooled testing. Our ordered rectangular-pooled testing proposed could take the advantage of their concept of zones, which will likely have different measured positivity rates. Hence, combining with our new approach would improve the testing efficiency.

Error handling. The response to viral testing, including SARS-Cov-2, is a function of the viral load, i.e., how much virus is present in the sample. Pooling dilutes the virus and hence can reduce the sensitivity. One of the major limitations of pooled testing is the exaggerated impact of false-negatives (FNs), because if a pool tests negative, all individuals in the pool are excluded from further testing. The impact is even larger in double-pooled testing since it requires both pools to test positive. One approach to addressing this, based on Kim et. al. (9), is to consider any occurrence of a row (column) that tests positive for which there is no corresponding column (row) intersection as a set requiring individual testing. Testing the unmatched rows/columns increases the cost slightly, but helps to reduce the impact of false-negatives. The work in (9) modeled the cost increases with testing errors, with a multitude of assumptions including a square design, constant probability, and multiple strong independence assumptions including that specificity and sensitivity are independent of the pool size. Because of the many assumptions needed, we have not tried to generalize the formula in (9) or include it in our cost estimations. However, in operational use, this error handling practice should be followed.

5 CONCLUSIONS AND FUTURE WORK

Pooled testing sensitivity is a function of both the dilution from multiple negative samples combined with the number of PCR cycles taken in amplification. Furthermore, public health professionals operating test facilities may want to limit the maximum pool size based on their actual test specificity/selectivity. Practical limitations such as the volume of pool collection equipment may also limit the pool sizes. Our free open-source software implementing the ordered rectangular-pooled testing allows user specified maximum row/column pool sizes to address such concerns.

While we have presented our work with respect to COVID19 testing, with added discussion of PCR-based techniques, the approach is a general model for two-stage pooled testing and can be applied to any problem where pooled testing with individual-testing confirmations are used. This includes standard disease (influenza and HIV) testing (13), experimental designs for object detection (14), and many other applications. With the asymmetric probabilities and double-pooled nature, it would be useful to extend the improved confidence bounds computed in (15) to this new model.

The proposed approach improves on the previous "optimal" double-pooled testing by relaxing the constraint of an assumed uniform distribution and makes use of per-item estimated probabilities. Given even rough estimates of the probability of positivity, the proposed ordered rectangular-double-pooled testing approach is easy to implement and significantly reduces the number of tests needed. The amount of gain increases as the population includes more negative samples, exactly the type of distribution expected as we scale testing to community monitoring rather than primarily diagnosis. One of the primary reasons for medical (and laboratory) errors is the mishandling or labeling of specimens, which could cause difficulties

with any pooled testing (16). Thus we are expanding our open-source software implementation of this approach to include tracking samples through the process. It starts with assigning them to pools, tracking test results, and identifying samples that need follow-up testing. The design of the software will be expanded to include such operational support functionality.

REFERENCES

1 .Dorfman R. The detection of defective members of large populations. *The Annals of Mathematical Statistics* **14** (1943) 436–440.

2 .de Wolff T, Pflüger D, Rehme M, Heuer J, Bittner MI. Evaluation of pool-based testing approaches to enable population-wide screening for covid-19. *arXiv preprint arXiv:2004.11851* (2020).

3 .Yelin I, Aharony N, Shaer Tamar E, Argoetti A, Messer E, Berenbaum D, et al. Evaluation of covid-19 rt-qpcr test in multi-sample pools. *Clinical Infectious Diseases* (2020).

4 .Wacharapluesadee S, Kaewpom T, Ampoot W, Ghai S, Khamhang W, Worachotsueptrakun K, et al. Evaluating the efficiency of specimen pooling for pcr-based detection of covid-19. *Journal of Medical Virology* (2020).

.[Dataset] Pooled sample testing and screening testing for covid-19. `https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/pooled-sample-testing-and-screening-testing-covid-19` (2020). [Last accessed Nov-27-2020].

5 .Lopez SM, Martin JM, Johnson M, Kurs-Lasky M, Horne WT, Marshall CW, et al. A method of processing nasopharyngeal swabs to enable multiple testing. *Pediatric research* **86** (2019) 651–654.

6 .Eberhardt JN, Breuckmann NP, Eberhardt CS. Multi-stage group testing improves efficiency of large-scale covid-19 screening. *Journal of Clinical Virology* (2020) 104382.

7 .Broder AZ, Kumar R. A note on double pooling tests. *arXiv preprint arXiv:2004.01684* (2020).

8 .Mézard M, Toninelli C. Group testing with random pools: Optimal two-stage algorithms. *IEEE Transactions on Information Theory* **57** (2011) 1736–1745.

9 .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* **63** (2007) 1152–1163.

10 .Phatarfod R, Sudbury A. The use of a square array scheme in blood testing. *Statistics in Medicine* **13** (1994) 2337–2343.

11 .Berger T, Mandell JW, Subrahmanya P. Maximally efficient two-stage screening. *Biometrics* **56** (2000) 833–840.

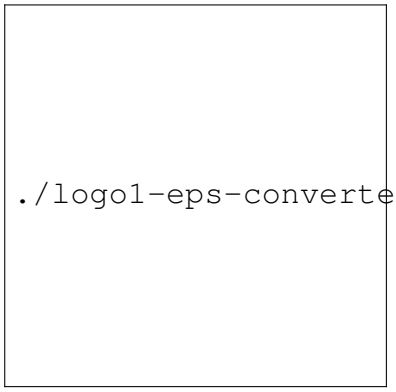
12 .Friedman E, Friedman J, Johnson S, Landsberg A. Transitioning out of the coronavirus lockdown: A framework for evaluating zone-based social distancing. *Frontiers in Public Health* **8** (2020) 266.

13 .Nguyen NT, Aprahamian H, Bish EK, Bish DR. A methodology for deriving the sensitivity of pooled testing, based on viral load progression and pooling dilution. *Journal of translational medicine* **17** (2019) 252.

14 .Anirudh R, Turaga P. Interactively test driving an object detector: Estimating performance on unlabeled data. *IEEE Winter Conference on Applications of Computer Vision* (IEEE) (2014), 175–182.

15 .Liu C, Liu A, Zhang B, Zhang Z. Improved confidence intervals of a small probability from pooled testing with misclassification. *Frontiers in public health* **1** (2013) 39.

16 .[Dataset] Rohde R. Covid-19 pool testing: Is it time to jump in? `https://asm.org/Articles/2020/July/COVID-19-Pool-Testing-Is-It-Time-to-Jump-In` (2020). [Amercian Society for Microbiology Online; accessed Aug-2020].



Supplementary Material

1 EXAMPLES

To help see the potential efficiency gains, we present a few examples to provide the reader with some intuition. The core idea comes from the fact that we have two conflicting objectives – rejecting negatives as much as possible, but also reducing ambiguity so we lower the number of individual tests needed. By first assuming two infection probabilities for two sub-populations, we start with two different pool sizes: the larger ones help in rejecting negative samples from the low-probability population, while the smaller ones reduce ambiguities by packing higher-probability samples into the same pool.

In practice, we will not be able to know the exact probabilities of infection. However, we can easily break the population down into two groups: Asymptomatic (A) and Symptomatic (S), wherein the general infection probabilities $p_A < p_S$. Generalizing this idea further, it would be natural to consider at least four groups where we can estimate p_i based on the questions asked during sample collection: Asymptomatic with unknown contacts (AU), Asymptomatic with known contacts (AK), Symptomatic with unknown contacts (SU), and Symptomatic with known contacts (SK). These four groups would, in general, have an increased probability of infection, $p_{AU} < p_{AK} < p_{SU} < p_{SK}$, which might be locally estimated based on recent results of testing for each group. One might further refine the estimate of the probability of an individual based on the amount of contact, the severity and the number of symptoms, or other factors. For simplicity, we start the analysis based on the two-group classification. With a bit of effort, one can derive formulas based on the general per-individual probabilities, given a set of samples, $s_i : i = 1 \ldots N$, each of which has a probability of infection p_i. As we shall see, our final approach uses individualized probabilities per pool optimization of the design.

The intuition behind our approach is that, assuming the sensitivity of a test is not a strong function of pool size, then a low probability of infection results in a larger optimal pool size, while a higher probability of infection requires a smaller pool size, potentially even individual testing. In our designs, the proportion and probability of the Asymptomatic (A) group directly impacts the number of columns. In contrast, the proportion and probability of the Symptomatic (S) group drives the number of rows. If the probabilities of A and S are assumed equal, it results in a standard square-double-pooled testing.

The second aspect of our improvement is optimizing row and column assignments to reduce individual testing. Because in double-pooled sampling, we individually test any sample that has both a positive row and positive column test, given r rows with $\check{r}$ positive row tests, and c columns with $\check{c}$ positive column tests, the number of individual tests needed is $\check{r} * \check{c}$. If there are $M < \min(r, c)$ positive subjects, a random

assignment will produce $\frac{M}{r} * \frac{M}{c}$ items on average that need individual testing. With a constant probability, this is optimized when $r = c$. However, with non-uniform probability, when we sort the data based on probability estimates, and fill the matrix in a row-major order, then the probability of rows testing positive increases with row number. Our goal is to reduce the number of rows expected to test positive, or $\overrightarrow{r}$, which is achieved through sorting.

Example 1. Let us return to the example in Figure 1 with a bit more detail. Recall there were Asymptomatic (A) and Symptomatic (S) groups where $p_A = .01, p_S = .2$ respectively. Among the 100 asymptomatic cases there is one positive, and within the 20 symptomatic cases there are four positive. The probability of infection for the overall mixed population is $.042$. For single-pooled testing, the optimal pool size for $p = .04$ is 5 (**?**), resulting in 24 pooled tests. Given the probabilities, out of these 24 tests, 5 will generally be positive, hence needing 25 (5×5) individual tests to disambiguate the results. This finally yields a total of 49 tests. For square-double-pooled testing with $p = .042$, the optimal pool size is 11 (**?**). In this example, it results in 11 rows and 11 columns (22 pooled tests), of which 5 rows and 5 columns will, in general, be positive, as shown in Figure 1b. This eventually results in the need for 25 individual tests, or a total of 47 tests.

Consider an ordered rectangular-pooled test with 6 rows and 20 columns. This produces five row-pools of size 20, each filled with samples from the asymptomatic group, and one row-pool also of size 20, filled with samples from the symptomatic group that has a high probability of infection, as shown in Figure 1a. In this case, 2 rows and 5 columns will test positive. With this ordered rectangular design, there will be 26 pooled tests (20 column-pools, 6 row-pools), followed by 10 (2×5) individual tests, for a total of 36 tests.

To see the impact of sorting on the design, consider if we used an 11x11 matrix but used ordered data. Then we would have 22 pooled tests, and if 3 rows and 5 columns test positive, we would need 15 individual tests to identify the 5 underlying cases. Thus, the expected total cost would be 37, just slightly higher. If, however, all of the positive instances from the symptomatic group happened to occur in the same row, then there would be only 2 rows and 5 columns test positive in the entire matrix, needing only 10 individual tests. Hence, the total cost would be 32. Thus, ordering is very important, and actual optimization requires a more detailed analysis of the expected costs.

Example 2. Now consider the same total population (120) and average probability of infection (.042), but this time assume there are 60 people with $p_A = .0166$ (again one positive case in the asymptomatic group), and 60 symptomatic people with $p_S = .0666$ (again four positive cases in the symptomatic group) which yields the same overall population positive rate of $.042$. The analysis for the single-pooled and square-double-pooled designs would still be the same, with the same number of tests, 49 and 47, respectively. In our ordered rectangular design with 6 rows and 20 columns, we would have 3 rows of samples from the asymptomatic group, of which 1 would be positive. For the samples from the symptomatic group, there would now generally be two, and at most three rows with positive testing response. Therefore, the overall would be at most 4 positive rows and 5 positive columns, needing 26 pooled tests and 20 individual tests for a total of 46 tests. Not as much savings, but still better than either of the other approaches. If, however, we used a 12-row by 10-column design, we would have 22 pooled test and 20 individual tests, for a total of 42 tests. Again, for simplicity, this example analysis ignores the full probabilities for which there will often be fewer rows/columns – the formulas derived below use the overall expected values. We do note that as the proportion and probability of positivity of the two groups gets closer, the optimal asymmetric design gets closer to the square-double-pooled design. If the same 12×10 design was used for the first example, it would be expected to require the same number of tests – errors in probability estimation reduce, but in general do not eliminate, the savings.

It is natural to ask what would happen if one processed the two sub-groups in two separate pools. In the first example, that would be a group of 100 and another one of 20. For the 100 cases in the asymptomatic group using double-pooled testing, it would require a 10×10 matrix, with one positive row and column, so no individual test would be needed. For the remaining 20 samples at $p_S = 0.2$, double-pooled testing would not be as effective as single-pooled or individual testing. With single-pooled testing at this positivity rate, the optimal pool size is three. Therefore, this requires 20 double-pooled tests for the asymptomatic group, and 7 single-pooled tests for the symptomatic group, of which, in general, 4 would be positive. This would require 12 (3×4) individual tests, for an overall total of 19 tests for the symptomatic cases (individual testing would require 20 individual tests). Therefore, separated pools would cost 39 tests, which is more efficient than the single- or standard double-pooled testing, but not as efficient as our ordered rectangular pooled testing. In the second example, the samples would be separated into two groups of 60. The first would form an 8×8 double-pool, of which one row and one column would be positive. The second group of the 8×8 double-pool would produce 4 positive rows and 4 positive columns, requiring 16 individual tests. Thus processing as two pools would require a total of 16+16+16 = 48 tests. This is about the same as the standard pooled testing, but less efficient than our ordered rectangular-pooled testing.

2 A SIMPLIFIED TWO-POPULATION MODEL

To provide some intuition where the main result formula comes, we derive the formula for a simpler model with two sub-populations, which we generalized for the main approximation.

We start by considering the standard single-pooled testing, given n samples in the pool with a uniform probability of infection p. First, the entire pool is tested (for a cost of 1), and then each item is tested (for a cost of n) if the result of the pool is positive. The pooled test is positive if every item is not negative, which happens with probability $(1-(1-p)^n)$. Therefore, the expected total cost for a single pool of size n is $(1+n(1-(1-p)^n))$.

To move toward the ordered rectangular-double-pooled testing, for simplicity, presume that there were sufficient samples from the two populations to fill any $r \times c$ matrix design. To simplify the analysis, assume the number of samples is n_1 and n_2 per sub-group, where $n_1 + n_2 < r \times c$, and both groups could fully fill their rows such that $r = r_1 + r_2$, with $r_1 = \frac{n_1}{c}, r_2 = r - \frac{n_1}{c}$. In other words, we assume that rows in r_i are filled with c items, all with a probability p_i of being positive. The expected number of positive samples is $M = p_1 * n_1 + p_2 * n_2$, where for simplicity we initially assumed $M < min(r, c)$. Therefore, on average, there are rows/columns that do not have any positive samples.

A column (row) is positive if all items in the column (row) are not negative. If an item's probability of being positive is p_1 in each of the first r_1 rows, p_2 in the remaining rows, and columns are independent and identically distributed, then the expected number of positive columns can be computed as $\overrightarrow{c} = c * \left(1-(1-p_1)^{\frac{n_1}{c}}(1-p_2)^{\frac{n_2}{c}}\right)$. In the first r_1 rows, the probability of a row being positive is $1-(1-p_1)^c$. Similarly, the remaining rows have a probability of $1-(1-p_2)^c$ being positive, resulting in the expected number of positive rows $\overrightarrow{r} = (\frac{n_1}{c} * (1-(1-p_1)^c) + (\frac{n_2}{c}) * (1-(1-p_2)^c))$. Combining the above, the overall expected total cost $TC(r, c)$ is

$$TC(r,c) \quad \approx \quad r + c + \overrightarrow{r} * \overrightarrow{c} \tag{S1}$$

where

$$\overrightarrow{r} = \frac{n_1}{c} * (1 - (1 - p_1)^c) + \frac{n_2}{c} * (1 - (1 - p_2)^c) \tag{S2}$$

$$\overrightarrow{c} = (c * \left(1 - (1 - p_1)^{\frac{n_1}{c}} (1 - p_2)^{\frac{n_2}{c}}\right) \tag{S3}$$

subject to $n_1 + n_2 = N; r * c >= N$ and $p_1 * n_1 + p_2 * n_2 < r + c$. The estimated per-item cost would be $\frac{TC(r,c)}{n_1+n_2}$.

To determine the optimal number of rows and columns, we can take partial derivatives of the equation above with respect to c and r, and set them to 0. The resulting functions are complex and cannot be solved algebraically; even if they could, the minimum point may not satisfy the required constraints. Alternatively, given that pooled testing is limited by the dilution effects and the pool size cannot exceed 64, one can exhaustively search the space for the globally optimal value, which is what we do in our code.