Some Simple Economics of Testing in a Pandemic: Efficiency of Pooled Testing Increases with Test Frequency

Ned Augenblick, Jonathan Kolstad, Ziad Obermeyer, Ao Wang
UC Berkeley

February 11, 2021

Abstract

Pooled testing increases efficiency by pooling individual samples and testing the combined sample, such that many individuals can be cleared with one negative test. This short paper demonstrates that repeated testing – as is needed in a pandemic – makes pooled testing particularly advantageous. First, repeated testing mechanically lowers the infection probability at the time of the next test by removing positives from the population. This effect alone means that increasing frequency by x times only increases expected tests by around $\sqrt{x}$. However, this calculation omits a further benefit of frequent testing: removing infections from the population lowers intra-group transmission, which lowers infection probability and generates further efficiency. For this reason, increasing testing frequency can paradoxically *reduce* total testing cost. Second, we show that pool size and efficiency increases with intra-pool risk correlation, which is expected given spread within natural testing pools (e.g., in workplaces, classrooms, etc). We conclude that frequent pooled testing using natural groupings is a cost-effective way to provide consistent testing of a population to suppress infection risk.

1 Introduction

The COVID-19 pandemic has generated a health and economic crisis not seen in more than a century. Opening businesses and schools is necessary to regain economic activity, but the potential public health costs are dramatic. One policy to circumvent this stark trade-off is to open the economy, while implementing *surveillance testing* that can quickly identify infected individuals – particularly those without symptoms – and prevent them from spreading the disease. Unfortunately, testing at this scale appears infeasible given the cost and capacity constraints. This paper makes a simple but essential point about these costs: when using pooling testing, frequent testing of correlated samples makes testing dramatically more efficient (and therefore less costly) than understood both by existing research and policy makers.

In pooled testing (Dorfman (1943)), multiple samples are combined, tested together using one test, and the entire pool is cleared given a negative test result. Pooling is an old concept and a large literature has emerged on optimal strategies (Dorfman (1943); Sobel and Groll (1959); Hwang (1975); Du et al. (2000); Saraniti (2006); Feng et al. (2010); Li et al. (2014); Aprahamian et al. (2018, 2019); Lipnowski and Ravid (2020)) and, more recently, others have discussed how it might be used to increase COVID test efficiency (Lakdawalla et al. (2020); Shental et al. (2020)). However, all of these papers focus on one-time testing of a set of samples with independent infection risk, which matches common use-cases such as screening donated blood for infectious disease (Cahoon-Young et al. (1989); Behets et al. (1990); Quinn et al. (2000); Dodd et al. (2002); Gaydos (2005); Hourfar et al. (2007)). These environmental assumptions are violated when dealing with a novel pandemic with rapid spread. In this case, people need to be tested multiple times and testing pools are likely formed from populations with correlated infection risk. How do these changes impact testing strategy?

We start with the well-known observation that pooled testing is more efficient when infection probability is lower, because the likelihood of a negative pooled test is increased. This observation has been used to conclude that pooled testing is not cost-effective for "high-prevalence" populations, such as health care workers or for people in areas experiencing an outbreak. While this statement is true for one-off testing, it does not hold when the population is tested repeatedly. As an extreme example, if a person in a high-prevalence area was just tested and determined to be negative, their probability of infection when tested an hour later is extremely low, simply because there is not much time to be infected between the tests. In other words, infection probability at the time of testing depends both on the flow rate of infection as well as the timing of testing.

We quantify the impact of testing frequency on infection probability and it's consequent impact on pooled-testing efficiency. For example, we show that given reasonable levels of independent risk, testing twice as often cuts the infection probability at the time of testing by (about) half, which lowers the expected number of tests at each testing round to about 70% of the original number. The savings are akin to a "quantity discount" of

30% in the cost of testing. Therefore, rather than requiring two times the numbers of tests, doubling frequency only increases costs by a factor of 1.4. More generally, we demonstrate that testing more frequently requires fewer tests than might be naively expected: increasing frequency by x times only uses about $\sqrt{x}$ as many tests, implying a quantity discount of $(1 - 1/\sqrt{x})\%$.

The benefits to frequency are even greater when there is intra-pool spread. In this case, testing more frequently has an additional benefit: by quickly removing infected individuals, infection spread is contained, future infection probabilities are lowered, and testing efficiency rises further. We show that in this case – somewhat paradoxically – the quantity discount is so great that more frequent testing can actually *reduce* the total number of tests.[1]

Finally, we show that intra-pool infection correlation increases the benefits of pooled testing even in a static testing environment. Intuitively, increased correlation in a pool with fixed individual risk lowers the likelihood of a positive pooled test result, which increases efficiency. Consequently, we conclude that pools should be formed from people who are likely to infect each other, such as those in a work or living space. This has a key logistical advantage: it implies that simple collection strategies – such as collecting sequential samples by walking down the hallway of a nursing home – can encode physical proximity and therefore capture correlations without sophisticated modeling.

To present transparent results, we consider a very stylized environment with a number of simplifications. While removing these constraints further complicates the problem and raises a number of important logistical questions, we do not believe that their inclusion changes our main insights. To pick one important example, we model a test with perfect sensitivity and specificity, but there is a natural concern that the sample dilution inherent in pooled testing leads to a loss of test sensitivity. However, the sensitivity loss of pooled testing given reasonable pool sizes has been shown to be negligible in other domains (Shipitsyna et al. (2007); McMahan et al. (2012)) and more recently shown to be similarly low for SARS-CoV-2 in pool sizes of 32 and 48 (Hogan et al. (2020); Yelin et al. (2020). Furthermore, even if there is a loss of sensitivity on a single test, this is counteracted by the large increase in overall sensitivity coming from running a larger number of tests given increased frequency.[2] Finally, if specificity is a concern, the past literature (Litvak et al. (1994); Aprahamian et al. (2019)) has clear methods to to optimize in the case of imperfect tests. There are multiple other issues, from pooling costs to regulatory barriers, but we believe the efficiency gains from frequent pooled testing are so high that addressing these issues is likely worth the cost.

Although we see this paper as noting a general insight of the relationship between pooled testing and testing

[1] As we show, these results do not rely on complex optimization of pool size or sophisticated cross-pooling or multi-stage pooled testing. Furthermore, the results are qualitatively similar when using a range of reasonable pool sizes.

[2] For example, if pooled testing leads the sensitivity to drop from 99% to 90% on a single test, sampling x times as frequently will increase overall sensitivity to $1 - (0.10)^x$ if errors are independent. Even with extremely correlation in the error – suppose the false negative rate for a pool given a previous false negative for that pool is 50% – pooled testing 4-5 times as frequently will recover the same false positive rate as individual testing.

frequency, it is useful to discuss the particular historical context in which the paper was written. The first paper draft of the paper was completed in June 2020, during the first wave of the COVID pandemic. At that point, testing supply was low and prices were high because laboratories were building up testing capacity in a relatively strict regulatory environment. By the time of the writing of this draft in February 2021, multiple organizations – such as Mirimus, Ginkgo, and the Broad – are offering frequent pooled testing at much cheaper prices than individual testing and multiple organizations with correlated risk – such as employers, cities, and school districts – are employing these tests. For example, in February 2021, Massachusetts implemented a policy of providing universal weekly pooled testing for all K-12 students and faculty and staff.[3] Nationally, the Rockerfeller Foundation has also called for use of frequent pooled testing as an essential aspect of school reopening (Aspinall (2020)).[4] The authors, based on the main insight of this paper, supported many of these policy initiatives and recommendations.

The paper proceeds as follows: Section 2 reviews a main finding in the pooled testing literature that efficiency rises as infection probability falls; Section 3 discusses the relationship between testing frequency and efficiency; Section 4 demonstrates how correlated infection leads to larger pool sizes and greater efficiency; and Section 5 concludes.[5]

2 Pooled Testing: Benefits Rise as Infection Probability Falls

2.1 Background on Pooled Testing

To understand the basic benefits of pooled testing, consider a simple example: 100 people, each with an independent likelihood of being positive of 1% and a test that (perfectly) determines if a sample is positive. The conventional approach of testing each person individually requires 100 tests. Suppose instead that the individuals' samples are combined into five equally-sized pools of 20 and then each of these combined samples are tested using one test. If any one of the 20 individuals in a combined sample is positive, everyone in that pool is individually tested, requiring 20 more tests (21 in total). The probability that this occurs is $1-(1-.01)^{20} \approx 18\%$. However, if no one in the pool is positive – which occurs with probability $\approx 82\%$ – no more testing is required. Because the majority of tests require no testing in the second case, the expected number of tests for this simple pooling method is only around 23, a significant improvement over the 100 tests required in the non-pooled method.

[3] See e.g. https://covidedtesting.com which provides a detailed description of the approach to pooled testing in schools and the experience in Massachusetts.

[4] This report notes that, even with vaccination, testing will continue to be a key policy lever. In wealthy countries, the potential for low vaccination take-up and much slower vaccination approval for children suggests that schools and possibly many employers will continue to need surveillance testing. Furthermore, many low income countries will not achieve large-scale vaccination for multiple years.

[5] In a previous version of the paper, we note that all pooled testing strategies rely on knowledge of infection probabilities, which are dynamic and difficult to estimate in a pandemic. We discuss how employing machine learning techniques on known data can create more accurate estimates and increase testing efficiency.

The approach is well studied with a large literature focused on improving the efficiency of pooled testing. These include using optimal pool size (e.g. in this example the optimal pool size of 10 would lower the expected number of tests to around 20), placing people into multiple pools (Phatarfod and Sudbury (1994)), and allowing for multiple stages of pooled testing (Sterrett (1957); Sobel and Groll (1959); Litvak et al. (1994); Kim et al. (2007); Aprahamian et al. (2018)). There are also methods to deal with complications, such as incorporating continuous outcomes (Wang et al. (2018)). Any of these modifications can be incorporated in our pooled testing strategy.

For clarity of exposition, we present results for simple two-stage "Dorfman" testing – in which every person in a positive pool is tested individually – to demonstrate that our conclusions are not driven by highly complex poolings and to make our calculations transparent.[6] As an example of this transparency, while the optimal pool size and associated efficiency formulas under Dorfman testing are complicated, components of low-order Taylor-Series approximations around infection probability $p = 0$ are very simple and accurate at the low probabilities needed for pooled testing.[7,8] Specifically, given a relatively-low infection probability p, the approximate optimal pool size is

$$g^* \approx \frac{1}{2} + \frac{1}{\sqrt{p}} \tag{1}$$

and the resultant approximate expected number of tests given a population of n people is

$$E[tests^*] \approx 2 \cdot \sqrt{p} \cdot n. \tag{2}$$

2.2 Infection Probability and Pooled Testing

For all of these different incarnations of pooled testing, the benefits of pooled testing rise as the infection probability falls in the population. Lower probabilities reduce the chance of a positive pooled test, thereby reducing the likelihood the entire pool must be retested individually. This is clear in Equation 2 as expected tests $2 \cdot \sqrt{p} \cdot n$ drop with infection probabilities. For example, if the probability drops from 1% to .1%, the optimal pool size rises and the number of tests falls from around 20 to 6.3. There is still a large gain if the pool size is fixed: expected tests drop from 23 to around 6.9 using a fixed pool size of 20. Similarly, if the probability rises from 1% to 10%, the expected number of tests using the optimal pool size rises to around 59 (or 93 given a fixed pool size of 20).

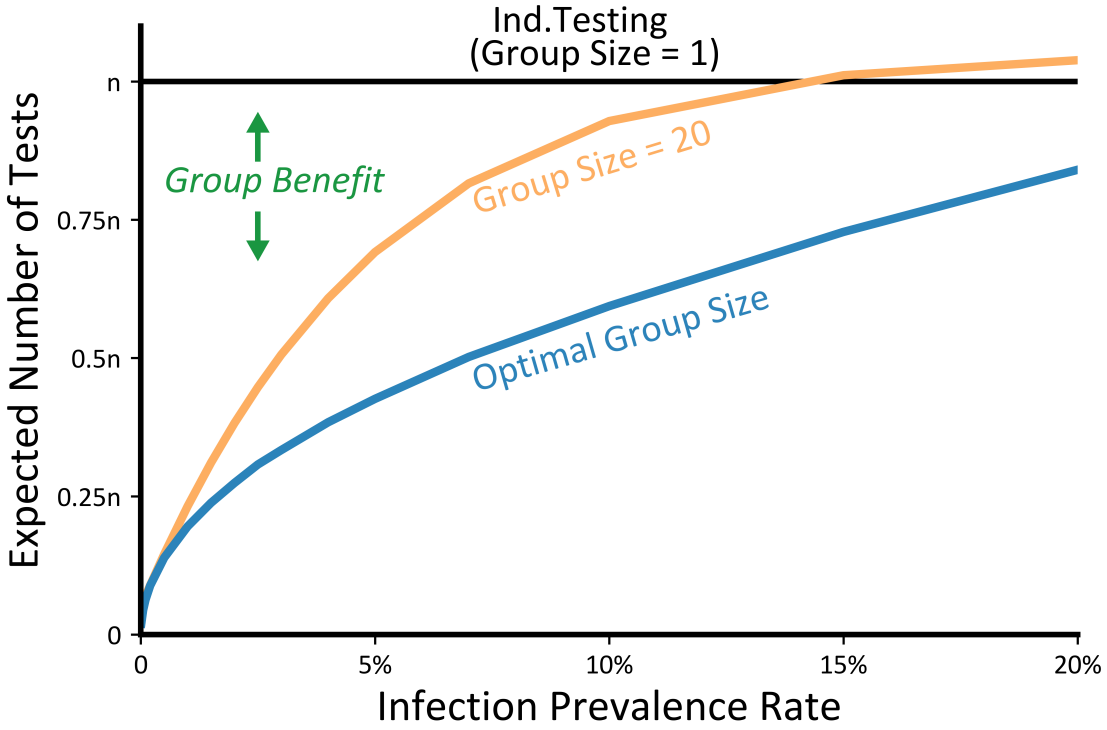
The full relationship is shown in Figure 1, which plots the expected number of tests in a population of n people

[6] In general, we advocate for these more sophisticated strategies when feasible as they further increase efficiency.

[7] The formula for the optimal pool size (disregarding rounding) is $g^* = {}^{2 \cdot W_0(-1/2\sqrt{-1/Ln(1-p)})}/_{Ln(1-p)}$ where $W_0(x)$ maps x to the principal solution for w in $x = we^w$. The expected number of tests is $(1 - e^{2 \cdot W_0(-1/2\sqrt{-1/Ln(1-p)})} + {}^{Ln(1-p)}/_{2 \cdot W_0(-1/2\sqrt{-1/Ln(1-p)})}) \cdot n$

[8] For the magnitude of infection probabilities we discuss in the paper, such as .1%, 1%, or 2%, the approximation of optimal pool size is within 0.3%, 0.1%, 0.01%, of the true optimal, respectively, and the approximation of the number of tests is within 3.1%, 2.3%, and 0.7% of the true number.

Figure 1: Efficiency of pooled testing rises with infection probability



Notes: This figure plots the expected number of tests (y-axis) from pooled testing given a population of n people as the population infection probability (x-axis) changes. The black flat line shows the number of tests from individual testing (equivalent to a pool size of 1), which always requires n tests regardless of infection probability. The results from using a pool size of 20 is orange, while the blue line represents the number of tests given the optimal pool size for a given probability. Finally, the green text notes that benefit from pooled testing is the distance between the black individual-testing line and those from pooled testing. For example, as noted in the text, using a pool size of 20 for a probability of 1% leads to $.23 \cdot n$ tests rather than n tests, while the optimal pool size (10) leads to $.20 \cdot n$ tests.

given different pool sizes and visually highlights the results based on (i) individual testing – which always leads to n tests, (ii) using pools of 20, and (iii) using optimal pooling given two stages. For simplicity, we construct these figures by assuming that n is large to remove rounding issues that arise from breaking n people into pools sizes that are not divisible by n.[9] There are large gains from pooled testing at many infection probabilities, though they are appreciably larger at lower probabilities.

[9]We note that this figure replicates many similar figures already in the literature going back to Dorfman (1943).

3 Increasing Test Frequency

3.1 Interaction Between Frequent Testing and Pooled Testing

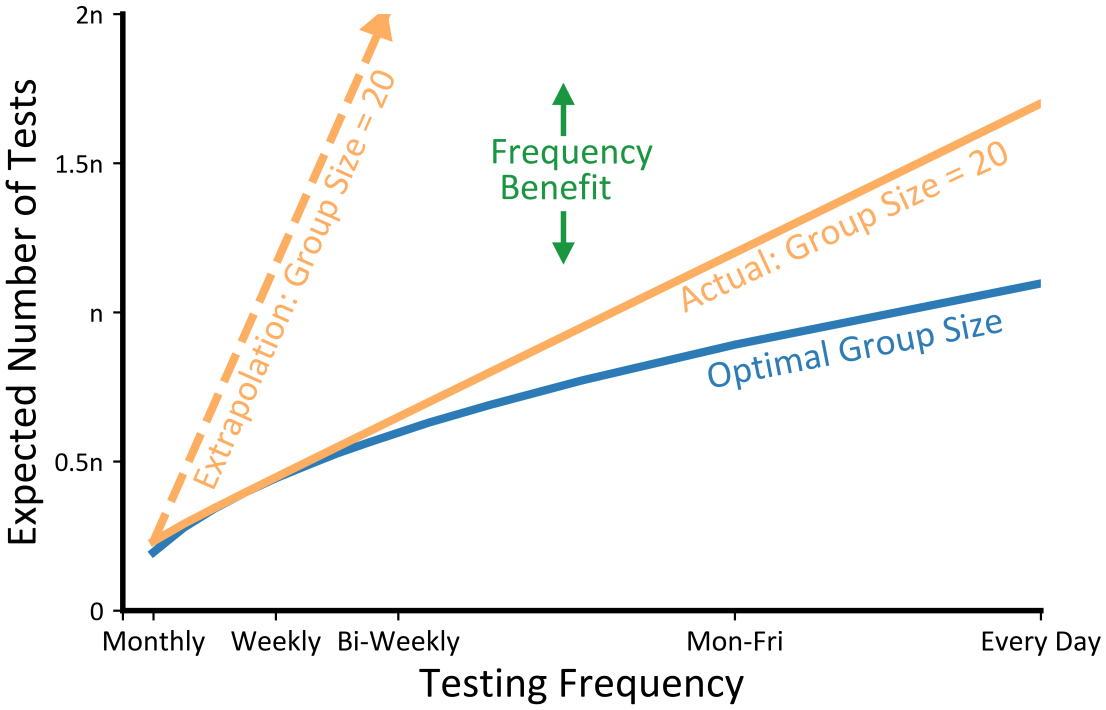
Our main insight is the important complementarity between pooled testing and testing frequency. Intuitively, the benefits of pooled testing rise as infection probability falls and frequent testing keeps this probability low at each testing period. Continuing with our example, suppose that 100 people have a 1% independent chance of being positive over the course a given time period. As discussed above, one could either sample everyone (requiring 100 tests), use pooled testing with a pool size of 20 (requiring ≈23 expected tests), or use pooled testing with an optimal pool size (requiring ≈20 expected tests).

Suppose instead that people are tested ten times as frequently. Testing individually at this frequency requires ten times the number of tests, for 1000 total tests. It is therefore natural think that pooled testing also requires ten times the number of tests, for more than 200 total tests. However, this estimation ignores the fact that testing ten times as frequently reduces the probability of infection at the point of each test (conditional on not being positive at previous test) from 1% to only around .1%. This drop in probability reduces the number of expected tests – given pools of 20 – to 6.9 at each of the ten testing points, such that the total number is only 69. That is, testing people 10 times as frequently only requires slightly more than three times the number of tests. Or, put in a different way, there is a quantity discount of around 65% by increasing frequency. The same conclusion holds using optimal pool sizes: the one-time pool test would require 20 expected tests, while testing ten times as frequently requires 6.3 tests at each testing point for a total of 63. The savings relative to the 1000 tests using individual testing are dramatic, with only approximately 6% of the total tests required.

Figure 2 represents this effect more generally for different levels of test frequency given an infection probability of 1% over the course of a month. Note that, at a frequency of a once a month, the numbers match those in Figure 1, which was based on one test given a probability of 1%. Unlike in Figure 1, we do not include the results for individual testing in this graph as testing individually everyday requires 20-30 times more tests than pooled testing, which renders the graph unreadable. The dotted orange line represents the naive (and incorrect) calculation for pooled testing by extrapolating the cost of testing multiple times by using the number of tests required for one test. That is, as above, one might naively think that testing x times using a pool size of 20 in a population of n would require $x \cdot .23 \cdot n$ tests given that testing once requires $.23 \cdot n$ tests. Pooled testing is in fact much cheaper due to the reduction in the probability of infection at the time of each time – the central contribution of this section. We therefore denote the savings between the extrapolation line and the actual requirements of pooled testing as the "frequency benefit."

The exact level of savings of the frequency benefit changes in a complicated way depending on the infection probability p given one test and the frequency x. However, using one component of the simple Taylor-Series

Figure 2: Efficiency of pooled testing rises with frequency



Notes: This graph presents the effect of testing frequency (x-axis) on the expected number of tests (y-axis), given infection probability for each individual in the population of 1% over a month. When the frequency is once a month, the points correspond to those in Figure 1 given probability of 1%: n for individual testing, $.23 \cdot n$ when using a pool size of 20 and $.20 \cdot n$ tests when using the optimal pool size. The dotted orange line represents the (incorrect) extrapolation that if a pool size of 20 leads to $.23 \cdot n$ tests when frequency is once a month, it should equal $x \cdot .23 \cdot n$ if frequency is x times a month. In reality, the expected tests are much lower, due to a quantity discount or "frequency benefit," highlighted by the green text. Finally, the blue line highlights tests given the optimally-chosen pool size.

approximation around $p = 0$ given the optimal pool size is again very accurate for reasonable infection rates[10] and makes the relationship clear:

$$E[tests^*|x] \approx 2 \cdot \sqrt{p} \cdot \sqrt{x} \cdot n. \tag{3}$$

Intuitively, testing at a frequency of x cuts the probability to around p/x, such that the expected tests at each testing time is around $2 \cdot \sqrt{p/x} \cdot n$, such that testing x times requires $2 \cdot \sqrt{p/x} \cdot x \cdot n$ total tests, which simplifies to Equation 3. Therefore, the expected cost of pooled testing x times as frequently is around $\sqrt{x}$ when using optimal-pool-sized two-stage pooled testing, and asymptotes to this exact amount as p falls to zero. In other words, the quantity discount of increased frequency is close to $(1 - 1/\sqrt{x})\%$. So, for example, pooled testing using optimally-sized pools every week (about 4 times a month) costs around $\sqrt{4} \approx 2$ times the number of tests from

[10] For example, even given $p = 5\%$, the approximation is within 1.3%, 0.6%, and 0.008% of the true number for x of 5, 10, and 100, respectively.

pooled testing every month, implying a quantity discount of 50%. Or, testing every day (around 30 times a month) costs about $\sqrt{30} = 5.5$ times the tests, implying a quantity discount of 82%.

3.2 Avoiding Exponential Spread Through Frequent Testing

The logic above ignores a major benefit of frequent testing: identifying infected people earlier and removing them from the population.[11] Beyond the obvious benefits, removing people from the testing population earlier stops them from infecting others, which reduces infection probabilities, and therefore increases the benefit of pooled testing. In the previous section, we shut down this channel by assuming that every person in the testing population had an independent probability of becoming infected. If the testing population includes people that interact, such as people who work or live in the same space, infections will instead be correlated.

Precisely modeling spread in a given population is challenging. The infection path is a complicated random variable that depends on exactly how and when different members of each particular testing population interact. Our goal is therefore not to make precise quantitative predictions but rather make a set of qualitative points based on the exponential-like infection growth common across virtually all models.[12] Consequently, we focus on the results of the simplest simulation that captures correlated exponential growth.

Specifically, we suppose that the 100 people continue to face a 1% chance of being infected over the course of a time period due to contact with someone outside of the test population. However, once a person is infected, this person can spread the infection to other people from within. We then chose an intra-pool instantaneous infection transmission rate such that if a person was infected at the start of the time period and was not removed from the population, the infected person would have a total independent probability s% over the time period of personally spreading the infection to each other person.[13] Now, consider the spread over the entire time period. With 37% probability, no one in the population is infected. However, with a 63% chance, a person is infected from the outside of the pool, setting off exponential-like growth within the rest of the test population (which ends only after testing and infected people are removed). Given this growth, for example, if s=1%, 5%, 10%, approximately 2, 8, and 26 people will be infected at the end of the time period. However, by testing ten times in the time period, infected people are removed more quickly and there is less time for a person to infect others, such that only around 1, 2, or 4 people are infected by the end of the time period.

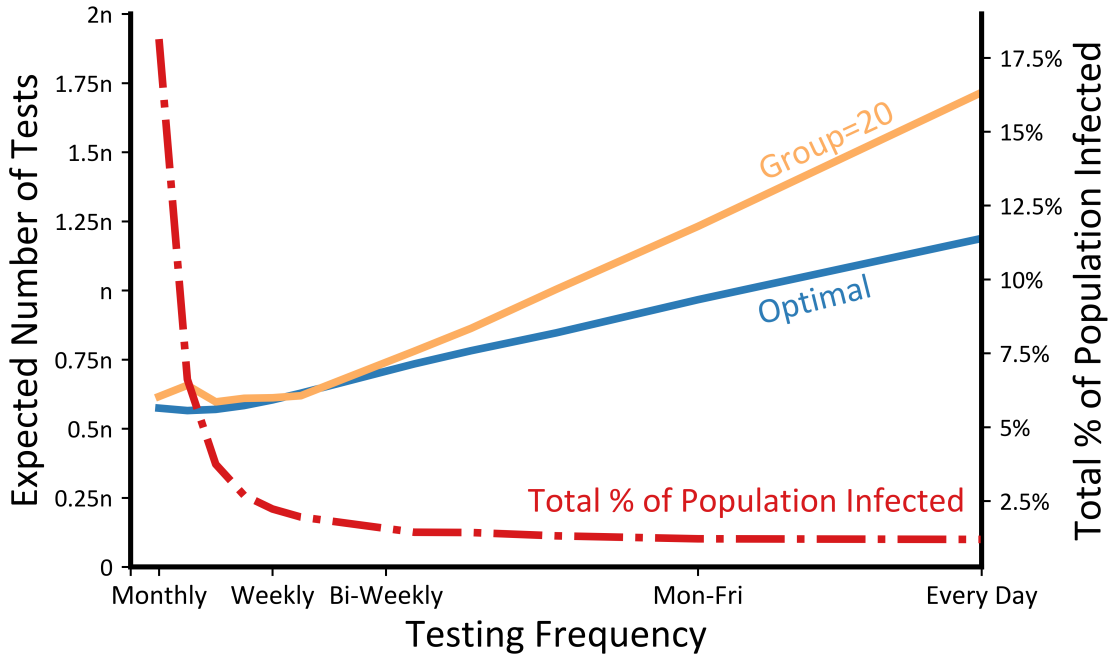
Not only are infections rates reduced, but the reduction in these rates reduces the relative cost of more

[11] Barak et al. (2020) notes a similar effect given a fixed budget of individual tests: it is more efficient to spread testing out over time because infected people are discovered earlier and removed.

[12] We consistently use the term "exponential-like" spread as we are considering smaller testing populations in which the infection rate slows as infected people are removed or become immune, such that the spread is actually logistic. However, we focus on situations in which the spread is caught early and therefore still close to exponential.

[13] For example, if s=5% and the time period is a month, then infected person i who has not been removed has a independent $1-(1-.05)^{(1/30)} \approx .17\%$ daily chance of infecting non-infected person j, $\approx .17\%$ daily chance of infecting non-infected person k, etc.

Figure 3: Increased frequency lowers infections with few additional tests given intra-pool spread



Notes: This graph presents the effect of testing frequency (x-axis) on the expected number of tests (y-axis 1) and final portion of the population that are infected (y-axis 2) given a model with intra-pool spread over the course of a month. As shown in red dot-dashed line of final infection proportions, increased frequency reduces the exponential-like spread because infected people are removed from the population. The number of expected tests required is shown for pool size 20 in orange and in blue for the optimal pool size. There is not much increase (and even an early decrease) in the total number of tests required because frequency increases leads to increased testing efficiency.

frequent pooled testing. For example, using a pool size of 20 and testing once over the time period leads to 23 tests in expectation given no spread. However, with spread given s=1%, 5%, 10%, this number rises to 30, 50, and 58, respectively. On the other hand, the expected number of tests when testing ten times as frequently does not grow at the same rate due to controlling the spread and increasing pooled testing efficiency. Specifically, the number of tests needed rises from 69 (with no spread) to 71, 84, and 115, respectively. That is, for example, at s=5%, the number of needed tests rises by 150% when testing once due to spread, but only rises by 22% when testing ten times as frequently.

These effects are shown in Figure 3. We plot the expected number of tests (left y-axis) and final portion of the population infected (right y-axis) for different testing frequencies assuming s=5%.[14] The number of infections rises in an exponential-like manner as frequency decreases and the infection is allowed to spread. The expected number of tests given different frequencies uses the same colors to represent pool sizes of 20 (orange) and optimal

[14]There is one minor difference between this figure and the example in the text. In the figure, we use a population of 120 rather than 100, as it allows for easier division into many pool sizes. This has a very slight impact as it slightly increases the time for the exponential spread to be curbed by increasing population immunity.

size (blue). Comparing Figures 2 and 3 is instructive. In Figure 2, we see a consistent increase in the tests required as the frequency of testing is increased. In Figure 3, though, the tests required are relatively flat and even decrease for early frequency increases. The difference is due to the fact that, with intra-pool infections, testing more frequently has the additional benefit of lowering the infection probabilities by containing infections, thereby increasing testing efficiency. For example, the quantity discount from a frequency of 2 is higher than 50% in the case of optimal pool sizes, such that the total cost from doubling the frequency actually falls.

3.3 Optimal Testing Frequency

The main benefit of increased frequency is reducing the exponential rise in infections. As shown in Figure 3, the marginal benefit from reduced infections due to increasing frequency is high at low frequencies and drops as frequency rises, eventually to zero. Interestingly, as shown in Figure 3, the number of tests can actually fall as frequency rises when starting at low frequencies. Therefore, for low frequencies, there is, in fact, no trade-off of raising frequency: it both reduces infections and reduces tests.

As testing frequency rises, the number of expected tests will inevitably rise, leading to a trade-off between marginal benefit and cost.[15] Consequently, at very high frequencies, there is an increased cost without a large benefit. The optimal frequency lies between these extremes, but depends on the value of reducing infections versus the cost of tests, which is an issue beyond the scope of this paper. However, our strong suspicion given our results and the economic costs of testing versus the economic (and human) costs of higher infections is that the optimal frequency is higher than the seemingly-common policy of testing on the order of monthly or only when an employee is symptomatic.[16]

4 Correlated Infection

In this Section, we discuss and isolate an additional factor implicitly included in the above example, which includes correlation between people whose samples are pooled together. We use a simple example to separate the insight that this correlation alone (i) is complementary with pooled testing in that it reduces the number of expected tests and (ii) leads to larger optimal pool sizes.

To first understand the benefit of correlation given pooled testing, it is useful to broadly outline the forces that determine the expected number of tests with simple two stage testing with a pool size of g and a large testing population n. In the first stage, a test will be run for every n/g pool, while in the second stage, every

[15] As an extreme example, if testing is so frequent that the infection probability at each test is effectively zero, then increasing the frequency by 1 will lead to an additional test for each pool without meaningfully reducing this probability at each testing period. This can be seen in Figure 3 for pool size of 20 where, at a frequency of around bi-weekly, the number of expected tests rises close to linearly with a slope of $1/20 = .05 \cdot n$.

[16] We also note a an important comparative static: it is more valuable to more frequently test a person who is more likely to catch and spread the infection (such as a person who meets with many other people versus a person who works alone in the back room).

n/g pool faces a probability q that at least one sample will be positive, such that all g people in the pool will need to be individually tested. Combining and simplifying these factors leads to a simple formula of the expected number of tests given a pool size: $n \cdot (1/g + q)$. As noted above, in the case of infections with independent probability p, $q = 1 - (1-p)^g$. However, as infections become more positively correlated, q falls for every pool size $g > 1$. For example, with two people in a pool whose infections have correlation r, q can be shown to be $1 - (1-p)^2 - r \cdot p \cdot (1-p)$. That is, when $r = 0$, we recover the original formula $1 - (1-p)^2$, while raising r linearly drops the probability until it is p when $r = 1$. Intuitively, the pool has a positive result if either person 1 *or* person 2 is infected, which – holding p constant – is less likely when infections are correlated and therefore more likely to occur simultaneously.

As an example of how q falls with more people and consequently reduces the number of tests, suppose that $p = 1\%$: when infections are uncorrelated, q is around 9.6%, 18.2%, 26.0%, and 33.1% given respective pool sizes 10, 20, 30, and 40, while q respectively drops to around 3.1%, 3.9%, 4.4%, and 4.8% when every person is pairwise-correlated with $r = 0.5$. Therefore, the respective expected number of tests given these pool sizes falls from $.196 \cdot n$, $.232 \cdot n$, $.294 \cdot n$, and $.356 \cdot n$ when uncorrelated to $.131 \cdot n$, $.089 \cdot n$, $.077 \cdot n$, and $.073 \cdot n$ when $r = 0.5$. First, note that the number of expected tests is universally lower at every pool size given correlation (and the savings are very significant). Second, note that while the pool size with the lowest number of expected tests given these potential pool sizes is 10 when there is no correlation, larger pool sizes are better given correlation. This statement is more general: higher correlation raises the optimal pool size. The intuition is that the marginal benefit of higher pool size (reducing the $1/g$ first-stage tests) is the same with or without correlation, but the marginal cost (increasing the probability of second stage testing) is reduced with higher correlation, thus leading to a higher optimum. As an example, while the optimal pool size given $p = 1\%$ is 10 given no correlation, the optimal pool sizes given r of 0, 0.2, 0,4, 0.6, 0.8 are 11, 22, 44, 107, and 385, respectively. Finally, note that when $r = 1$, the optimal pool size is unbounded, because adding an extra person to a pool adds benefit by reducing first-stage tests, but has no cost because the probability of a positive in the pool remains constant at p. Obviously, extremely large pool sizes are potentially technologically infeasible, but the intuition remains that correlation within testing pools should raise pool size.

5 Conclusions

This paper shows that pooled testing is particularly efficient when frequently performed on pools with correlated risk (e.g., in workplaces or schools). Our key insight is that repeated testing reduces the infection probability at the time of each test and – since pooled testing is more efficient given lower probabilities – increases the efficiency of pooled testing. Therefore, contrary to a commonly stated rule, pooled testing is appropriate and cost-effective

even for high-risk populations, as long as the frequency of testing rises in relation to this risk.[17] While a formal cost-effectiveness analysis is beyond the scope of this paper, we estimate that bi-weekly pooled testing could be offered for between $3-5 per person per day using the insights from this paper. In practice, frequent pooled testing is increasingly being adopted at scale, such as the state of Massachusetts where all K-12 students are offered weekly pooled testing. The cost of testing in those programs validates these estimates and demonstrates feasibility of these strategies.

[17] This statement is not true for the initial round of testing: when testing starts in a risk-risk population, infection probabilities are likely to be high. In this case, it is cost effective to perform a round of individual tests before switching to pooled testing.

References

Aprahamian, Hrayer, Douglas R Bish, and Ebru K Bish, "Optimal risk-based group testing," *Management Science*, 2019, *65* (9), 4365–4384. 1, 2

_ , Ebru K Bish, and Douglas R Bish, "Adaptive risk-based pooling in public health screening," *IISE Transactions*, 2018, *50* (9), 753–766. 1, 4

Barak, Boaz, Mor Nitzan, Neta Ravid Tannenbaum, and Janni Yuval, "Optimizing testing policies for detecting COVID-19 outbreaks," 2020. https://www.boazbarak.org/Papers/COVID19_arxiv.pdf. 8

Behets, Frieda, Stefano Bertozzi, Mwamba Kasali, Mwandagalirwa Kashamuka, L Atikala, Christopher Brown, Robert W Ryder, and Thomas C Quinn, "Successful use of pooled sera to determine HIV-1 seroprevalence in Zaire with development of cost-efficiency models.," *AIDS (London, England)*, 1990, *4* (8), 737–741. 1

Cahoon-Young, B, A Chandler, T Livermore, J Gaudino, and R Benjamin, "Sensitivity and specificity of pooled versus individual sera in a human immunodeficiency virus antibody prevalence study.," *Journal of Clinical Microbiology*, 1989, *27* (8), 1893–1895. 1

Dodd, RY, EP Notari IV, and SL Stramer, "Current prevalence and incidence of infectious disease markers and estimated window-period risk in the American Red Cross blood donor population," *Transfusion*, 2002, *42* (8), 975–979. 1

Dorfman, Robert, "The detection of defective members of large populations," *The Annals of Mathematical Statistics*, 1943, *14* (4), 436–440. 1, 5

Du, Dingzhu, Frank K Hwang, and Frank Hwang, *Combinatorial group testing and its applications*, Vol. 12, World Scientific, 2000. 1

et al. Aspinall, "How did covid-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data," Technical Report, Rockerfeller Foundation 2020. 3

Feng, Jiejian, Liming Liu, and Mahmut Parlar, "An efficient dynamic optimization method for sequential identification of group-testable items," *IIE Transactions*, 2010, *43* (2), 69–83. 1

Gaydos, Charlotte A, "Nucleic acid amplification tests for gonorrhea and chlamydia: practice and applications," *Infectious Disease Clinics*, 2005, *19* (2), 367–386. 1

Hogan, Catherine A, Malaya K Sahoo, and Benjamin A Pinsky, "Sample pooling as a strategy to detect community transmission of SARS-CoV-2," *Jama*, 2020, *323* (19), 1967–1969. 2

Hourfar, Michael Kai, Anna Themann, Markus Eickmann, Pilaipan Puthavathana, Thomas Laue, Erhard Seifried, and Michael Schmidt, "Blood screening for influenza," *Emerging infectious diseases*, 2007, *13* (7), 1081. 1

Hwang, FK, "A generalized binomial group testing problem," *Journal of the American Statistical Association*, 1975, *70* (352), 923–926. 1

Kim, Hae-Young, Michael G Hudgens, Jonathan M Dreyfuss, Daniel J Westreich, and Christopher D Pilcher, "Comparison of group testing algorithms for case identification in the presence of test error," *Biometrics*, 2007, *63* (4), 1152–1163. 4

Lakdawalla, Darius, Emmet Keeler, Dana Goldman, and Erin Trish, "Getting Americans back to work (and school) with pooled testing," *USC Schaeffer Center White Paper*, 2020. 1

Li, Tongxin, Chun Lam Chan, Wenhao Huang, Tarik Kaced, and Sidharth Jaggi, "Group testing with prior statistics," in "2014 IEEE International Symposium on Information Theory" IEEE 2014, pp. 2346–2350. 1

Lipnowski, Elliot and Doron Ravid, "Pooled Testing for Quarantine Decisions," Working Paper 2020-85, Becker Friedman Institute 2020. 1

Litvak, Eugene, Xin M Tu, and Marcello Pagano, "Screening for the presence of a disease by pooling sera samples," *Journal of the American Statistical Association*, 1994, *89* (426), 424–434. 2, 4

McMahan, Christopher S, Joshua M Tebbs, and Christopher R Bilder, "Informative dorfman screening," *Biometrics*, 2012, *68* (1), 287–296. 2

Phatarfod, RM and Aidan Sudbury, "The use of a square array scheme in blood testing," *Statistics in Medicine*, 1994, *13* (22), 2337–2343. 4

Quinn, Thomas C, Ron Brookmeyer, Richard Kline, Mary Shepherd, Ramesh Paranjape, Sanjay Mehendale, Deepak A Gadkari, and Robert Bollinger, "Feasibility of pooling sera for HIV-1 viral RNA to diagnose acute primary HIV-1 infection and estimate HIV incidence," *Aids*, 2000, *14* (17), 2751–2757. 1

Saraniti, Brett A, "Optimal pooled testing," *Health care management science*, 2006, *9* (2), 143–149. 1

Shental, Noam, Shlomia Levy, Shosh Skorniakov, Vered Wuvshet, Yonat Shemer-Avni, Angel Porgador, and Tomer Hertz, "Efficient high throughput SARS-CoV-2 testing to detect asymptomatic carriers," *medRxiv*, 2020. 1

Shipitsyna, Elena, Kira Shalepo, Alevtina Savicheva, Magnus Unemo, and Marius Domeika, "Pooling samples: the key to sensitive, specific and cost-effective genetic diagnosis of Chlamydia trachomatis in low-resource countries," *Acta dermato-venereologica*, 2007, *87* (2), 140–143. 2

Sobel, Milton and Phyllis A Groll, "Group testing to eliminate efficiently all defectives in a binomial sample," *Bell System Technical Journal*, 1959, *38* (5), 1179–1252. 1, 4

Sterrett, Andrew, "On the detection of defective members of large populations," *The Annals of Mathematical Statistics*, 1957, *28* (4), 1033–1036. 4

Wang, Dewei, Christopher S McMahan, Joshua M Tebbs, and Christopher R Bilder, "Group testing case identification with biomarker information," *Computational statistics & data analysis*, 2018, *122*, 156–166. 4

Yelin, Idan, Noga Aharony, Einat Shaer-Tamar, Amir Argoetti, Esther Messer, Dina Berenbaum, Einat Shafran, Areen Kuzli, Nagam Gandali, Tamar Hashimshony et al., "Evaluation of COVID-19 RT-qPCR test in multi-sample pools," *MedRxiv*, 2020. 2