

A Theory for Covid-19 Testing To Save Both Resources and Time

Chihjen Lee, MD, MPH, Allison Lee, BA, MAT, Lena Wang, MD

Acknowledgement: Special thanks to Christian Seif, BS, for his dedication to proofreading, editing, and revising of this manuscript.

Abstract

In Los Angeles, at one point, the Covid-19 testing positivity rate was 6.25%, or one in sixteen. This translates to, on average, one in sixteen specimens testing positive and the vast majority testing negative. Usually, we run sixteen tests on sixteen specimens to identify the positive one(s). This process can be time consuming and expensive. Since a group of negative specimens pooled together for testing will produce a negative result, one single test could potentially eliminate many specimens. Only when the pooled specimen tests positive do we need further testing to identify the positive one(s). Based on this concept, we designed a strategy that will identify the positive specimen(s) efficiently. Assuming one in sixteen specimens is positive, we find that only four tests are needed. Furthermore, we can run them simultaneously, saving both resources and time. Although, in the real world, we cannot make the same assumption of only one positive specimen, the same strategy works with slight modification and proves to be much more efficient than the conventional testing. Our strategy returns an answer 48% of the time in four tests and one time cycle. Overall, the average number of tests is seven or eight depending on the follow-up testing, and the average time cycle is about one and a half.

Keywords: Covid, Covid testing, sample pooling, resources, time, binary system, probability, positivity rate

Introduction

Suppose Covid testing positivity rate is 6.25%. That is, on average, one specimen would test positive and the other fifteen would test negative. Oftentimes, we find most of the Covid tests are negative, for instance, when we run screening tests for asymptomatic patients, workers, or students. Therefore, more efficient testing strategies were developed. One strategy was to pool specimens together [1,2]. Since pooling negative specimens together produces a negative test result, one single test would potentially eliminate many specimens. Thus, we present a theoretical approach to the Covid testing strategy, using sixteen specimens as an example. We will show how to pool specimens together, run the testing on different groups simultaneously, and identify the positive one(s) with binary encoding and decoding. This approach can be applied to different Covid testing positivity rates and any number of specimens.

Method

Suppose there are sixteen specimens, and all of them have enough sample for more than one test. We first present the strategy for a hypothetical scenario, then the strategy for the real world.

- 1) Assuming there is one and only one Covid positive specimen.

We pool eight specimens together to form a single new specimen for testing. Four groups will be formed according to the following strategy (Table 1).

Specimen No.	Group 1	Group 2	Group 3	Group 4
0				
1				█
2			█	
3			█	█
4		█		
5		█	█	█
6		█	█	
7		█	█	█
8	█			
9	█			█
10	█		█	
11	█		█	█
12	█	█		
13	█	█		█
14	█		█	
15	█		█	█

Table 1. Four groups of new specimens are formed, as marked in blue. The specimens in each group are listed below:

Group 1: 8 to 15

Group 2: 4 to 7, 12 to 15

Group 3: 2, 3, 6, 7, 10, 11, 14, 15

Group 4: 1, 3, 5, 7, 9, 11, 13, 15

All four pooled specimens are then submitted for PCR Covid testing simultaneously. A test would turn out positive if any one of the eight initial specimens is positive; otherwise, if all of them are negative, the result will be negative. Each group will either test positive or negative. The combination of positive and negative results will then identify which patient is Covid positive. Two scenarios are given below as examples:

In the first scenario, suppose group 1 is positive, group 2 negative, group 3 negative and group 4 positive (Table 2). Let positive results be given a value 1 and negative be given 0. Describe the overall test results in 0 and 1 combination from left to right, with left-most being group 1 and right-most group 4. Then, we can identify the Covid positive patient to be patient number 1001, in binary, i.e., patient number 9 in decimal. The conversion from binary to decimal is computed as follows:

$$1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 9$$

Scenario 1: Patient no.	Positive Group 1	Negative Group 2	Negative Group 3	Positive Group 4
0				
1				Positive
2			Negative	
3			Negative	Positive
4		Negative		
5		Negative		Positive
6		Negative	Negative	
7		Negative	Negative	Positive
8	Positive			
9	Positive			Positive
10			Negative	
11			Negative	Positive
12		Negative		
13		Negative		Positive
14		Negative	Negative	
15		Negative	Negative	Positive

Table 2. Positive (in red) and negative (in green) combinations identify the Covid patient. Of note, the pooled specimen can test either positive or negative. Under the assumption that one and only one patient is Covid positive, we know if the pooled specimen tests positive, then the second half of untested specimens must be negative and vice versa. Because group 1 tested positive in this scenario, we know that the positive specimen is present in specimens 8 to 15. Since group 2 tested negative, we can rule out specimens 12 to 15. Similarly, we further rule out specimens 10 and 11 from group 3 testing. Now only specimens 8 and 9 are left. Since group 4 tested positive, we know that specimen 9 is positive.

In the second scenario, suppose group 1 is negative, group 2 positive, group 3 negative and group 4 positive (shown in table 3). Again, let positive results be given a value 1 and negative be given 0. Describe the overall test results in 0 or 1 combinations from left to right, with left-most being group 1 and right-most group 4. Then, we can identify the Covid positive patient to be patient number 0101, in binary, i.e., patient number 5 in decimal. The conversion from binary to decimal is computed as follows:

$$0 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 5$$

Scenario 2: Patient no.	Negative Group 1	Positive Group 2	Negative Group 3	Positive Group 4
0				
1				Positive
2			Negative	
3			Negative	Positive
4		Positive		
5		Positive	Negative	Positive
6			Negative	
7		Positive	Negative	Positive
8	Negative			
9	Negative			Positive
10	Negative		Negative	
11	Negative		Negative	Positive
12	Negative	Positive		
13	Negative			Positive
14	Negative		Negative	
15	Negative		Negative	Positive

Table 3. The test results of scenario 2. Negative group 1 testing rules out specimens 8 to 15 and rules in specimens 0 to 7; positive group 2 testing rules in specimens 4 to 7; negative group 3 testing rules in specimens 4 to 5; positive group 4 testing rules in specimen 5.

As we can see from the above examples, the combinations of test results in 0 or 1 actually encode the positive specimen in binary. With this strategy, we only need to run four PCR tests on four pooled specimens instead of the conventional sixteen PCR tests. In addition, since we can run four PCR tests simultaneously, we can save both resources and time.

2) Without the knowledge of how many specimens are positive

In real world scenario, we cannot assume there is only one positive covid specimen. However, similar strategy could apply to the grouping and testing. In this new strategy, we still pool 8 specimens together, but we evaluate only 15 patients at a time, patient 1 to 15 (Table 4). We don't have patient number 0 this time, because, in the first strategy, we never test patient number 0, so we never know if he/she is positive or negative.

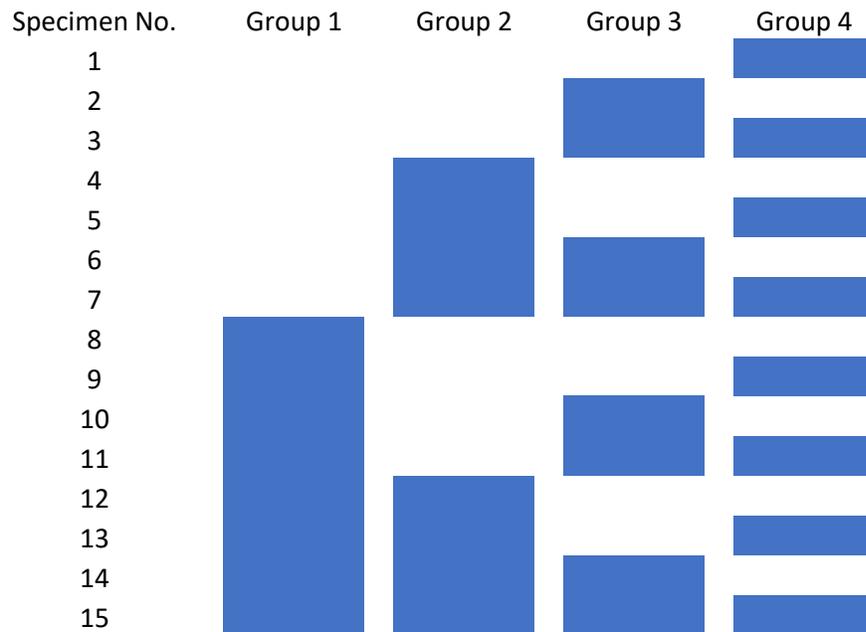


Table 4. Four groups of new specimens are formed, as marked in blue. The specimens in each group are listed below:

Group 1: 8 to 15

Group 2: 4 to 7, 12 to 15

Group 3: 2, 3, 6, 7, 10, 11, 14, 15

Group 4: 1, 3, 5, 7, 9, 11, 13, 15

With this strategy, there are five test results from which we can draw a final conclusion. Let positive be 1 and negative be 0. (1) When the result is 0000, all patients are Covid negative. (2) When the result is 0001 (table 5), patient number 1 is positive. (3) When the result is 0010, patient number 2 is positive. (4) When the result is 0100, patient number 4 is positive. (5) When the result is 1000, patient number 8 is positive. Test result of 0001 is illustrated in the following table.

Test Result 2: Patient no.	Negative Group1	Negative Group2	Negative Group3	Positive Group4
1				Positive
2			Negative	
3		Negative	Negative	Positive
4		Negative		
5		Negative	Negative	Positive
6			Negative	
7		Negative	Negative	Positive
8	Negative			
9				Positive
10			Negative	
11			Negative	Positive
12		Negative		
13			Negative	Positive
14			Negative	
15			Negative	Positive

Table 5. An example of one positive. Three negatives rule out specimens 2 to 15. The positive rules in specimen 1.

When the test result has two positives and two negatives, we can rule out 12 specimens and we might need up to 3 additional tests to find the positive specimen(s) (Table 6). This way, the total number of tests is $4 + 3 = 7$, which is still much less than the conventional 15 tests.

Two Positives: Patient no.	Negative Group1	Negative Group2	Positive Group3	Positive Group4
1				Positive
2			Positive	
3		Negative	Positive	Positive
4		Negative		
5		Negative	Positive	Positive
6			Positive	
7		Negative	Positive	Positive
8	Negative			
9				Positive
10			Positive	
11			Positive	Positive
12		Negative		
13			Positive	Positive
14			Positive	
15			Positive	Positive

Table 6. An example of two positive specimens. Two negatives rule out specimens 4 to 15. Specimens 1 to 3 need further testing. If specimen 3 tests negative, then we know both specimens 1 and 2 are positive. Otherwise, further testing is needed.

When the test result has three positives and one negative (Table 7), we can rule out 8 patients and we might need up to 7 additional tests. This way, the total number of tests is $4 + 7 = 11$, still less than the conventional 15 tests.

Three Positives: Patient no.	Negative Group1	Positive Group2	Positive Group3	Positive Group4
1				█
2			█	
3			█	█
4		█		
5		█		█
6		█	█	
7		█	█	█
8	█			
9	█			█
10	█		█	
11	█		█	█
12	█	█		
13	█	█		█
14	█	█	█	
15	█	█	█	█

Table 7. An example of three positives. The negative test rules out specimens 8 to 15. Specimens 1 to 7 need more testing.

When the test results are all positive, no conclusion can be drawn. All positive results suggest specimen 1111, or specimen 15, could be the culprit. However, many other possibilities would produce the same result. For example, positive specimens 13 and 14 will cause all pooled specimens to be positive. The most efficient strategy at this point is yet to be determined.

Discussion

Algorithms saving only resources. The Covid testing positivity rate varies from time to time and from city to city. The strategy presented here was developed for a positivity rate of 6.25%, i.e., one in sixteen. With the assumption of only one positive specimen out of sixteen, some other strategies save only resources [1]. They pool eight specimens together. If eight together is positive, then test four. If four

together is positive, then test two. If two is positive, then test one. During the process, a negative result will lead to the testing of half of the untested half. This way, four tests are needed to find the positive specimen. However, this strategy, although it saves resources, requires consecutive testing and therefore will spend four time-cycles. In contrast, our strategy potentially needs only one or two-time cycles and thus saves both resources and time.

Binary encoding. The four groups in our strategy can be formed using binary codes. In binary, sixteen specimens will run from 0000 to 1111, (or from 0001 to 1111 in the case of fifteen specimens). The first group, 8 to 15 group, is made up of all specimens whose fourth digit is 1, i.e., in the form of 1xxx, x being 0 or 1. The second group is made up of all specimens whose third digit is 1, i.e., in the form of x1xx, x being 0 or 1. So on, so forth. The beauty of the strategy is that the combinations of positive and negative test results in binary directly encodes the Covid positive patient's number as shown above.

Probability. In the real-world scenario, the probabilities of getting zero, one, two, three and four positive group-test results are 37.98%, 10.13%, 18.30%, 19.78% and 13.81% respectively. (See supplement 1, 2 and 3.) Therefore, our strategy returns an answer 48%, ($37.98\% + 10.13\% \approx 48\%$), of the time, after four tests and one time cycle. 18.3% of the time, up to three additional tests are needed. 19.78% of the time, up to seven additional tests are needed. Only 13.81% of the time does the testing result in 1111, and the best testing strategy from here is yet to be determined, needing possibly up to additional 15 tests. Overall, the average number of tests is **8**, by weighted averaging as shown below:

$$0.3798 \cdot 4 + 0.1013 \cdot 4 + 0.1830 \cdot 7 + 0.1978 \cdot 11 + 0.1381 \cdot 19 = 8.0051$$

The overall number of time cycles is 1.5.

$$0.3798 \cdot 1 + 0.1013 \cdot 1 + 0.1830 \cdot 2 + 0.1978 \cdot 2 + 0.1381 \cdot 2 = 1.5189$$

Therefore, in real-life situation, our strategy remains valid and efficient

Although there exist more sophisticated strategies to decipher the results of two, three or four positives, the additional saving of resources and time is limited. The average number of tests saved would be 1, dropping from 8 to 7. (See supplement 3.) For practical purpose, the easiest way is to test each of undetermined specimens.

Extending to 2^n patients. At different testing positive rates, this strategy can be modified. However, the same principle applies. Suppose we have 2^n patients. Then n groups are formed according to the following n-digit binary notations: 1xx...x, x1x...x, xx1...x, etc. When the number of specimens is between 2^n and 2^{n-1} , n groups are still formed with similar binary strategy. For example, twenty patients will be grouped as follows: 1xxxx, x1xxx, xx1xx, xxx1x, xxxx1, where the first group consists of five patients: patient 16, 17, 18, 19 and 20 or, in binary, patient 10000, 10001, 10010, 10011 and 10100.

Different positivity rates. At different positivity rates, the size of the pooled specimen needs to be modified accordingly. For instance, when the positive rate is 12.5% or one in eight, the best strategy is to pool seven specimens together. Similarly, when the positive rate is 3.125% or one in thirty-two, the best strategy is to pool 31 specimens together.

Limitations. Specimens are diluted in the process of pooling. A pooled specimen may contain less genetic material for PCR testing and, therefore, return a false negative result. Clinical validation is needed. For this strategy to work, it may require a test with higher sensitivity. Another limitation is that not all tests return a definite positive or negative result. When the result is intermediate, or indeterminate, more tests are needed [2].

Conclusion

Since most Covid tests are negative, it is wasteful of resources and time to test all specimens. Our strategy allows for pooling specimens together and testing different groups of pooled specimens simultaneously. Our strategy is applicable to different stages of the pandemic when the testing positivity rates are different. The binary encoding and decoding process during the testing is convenient and intuitive. In a hypothetical situation where only one in sixteen specimens is positive, our strategy needs only four tests and one time cycle to identify the positive one. For real world application, our strategy will return an answer 48% of the time in four tests and one time cycle (, given the positivity rate of 6.25%). Overall, the average number of tests is seven or eight and the average time cycle is around one and a half.

References

1. News. The mathematical strategy that could transform coronavirus testing. Nature 583, 504-505 (2020). DOI: <https://doi.org/10.1038/d41586-020-02053-6>

[The mathematical strategy that could transform coronavirus testing \(nature.com\)](#)

2. [Pooled Sample Testing and Screening Testing for COVID-19 | FDA](#)

Declarations:

1. There is no funding.
2. There is no conflict of interests.

Supplements (1): Probabilities

The probabilities of various test results are shown below.

1. Zero positive: 0000

$$\left(\frac{15}{16}\right)^{15} \approx \mathbf{0.3798}$$

2. One positive, four permutations: 0001, 0010, 0100, 1000

$$\left(\frac{15}{16}\right)^{14} \left(\frac{1}{16}\right) \times 4 \approx \mathbf{0.1013}$$

0001

Negative

Negative

Negative

Positive

Patient no.	Group1	Group2	Group3	Group4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

Table 8. Using 0001 as an example. The probability of specimen number 1 being positive is $\frac{1}{16}$; the probability of specimen numbers 2 to 15 being negative is $\left(\frac{15}{16}\right)^{14}$. Thus, the probability of 0001 test result is $\left(\frac{15}{16}\right)^{14} \left(\frac{1}{16}\right)$. There are four permutations, therefore the probability of one positive is $\left(\frac{15}{16}\right)^{14} \left(\frac{1}{16}\right) \times 4$.

3. Two positives, six permutations: 0011, 0101, 1001, 0110, 1010, 1100

The probability of two positives is obtained by brute force (Table 11): **0.1830**

Shown below are two cases of test result 0011.

0011 Patient no.	Negative Group1	Negative Group2	Positive Group3	Positive Group4
1	?	?	?	?
2	?	?	?	?
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

Table 9. Specimen number 3 is positive. Positive number 3 will render the test result 0011, regardless of the status of specimens 1 and 2.

0011 Patient no.	Negative Group1	Negative Group2	Positive Group3	Positive Group4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

Table 10. Specimen number 3 is negative. Positive specimens 1 and 2 will also render the test result 0011.

4. Three positives, four permutations: 0111, 1011, 1101, 1110
By brute force (Table 11), the probability is **0.1978** ...

5. Four positives: 1111

By brute force (Table 11), the probability is **0.1381** ...

Of note, the sum of all probabilities is 1:
 $0.3798 + 0.1013 + 0.1830 + 0.1978 + 0.1381 = 1$

Supplement (2): Brute force approach to probabilities

Patient Number															Group							
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	sum	Probability	1	2	3	4	sum	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3798	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0.0253	0	0	0	1	1	
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0.0253	0	0	1	0	1	
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0.0253	0	1	0	0	1	
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0.0253	1	0	0	0	1	
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0.0253	0	0	1	1	2	
.
0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	0.0017	0	0	1	1	2	
.
0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	3	0.0001	0	0	1	1	2	
.
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0.0253	0	1	1	1	3	
.
0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2	0.0017	0	1	1	1	3	
.
0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	3	0.0001	0	1	1	1	3	
.
0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	4	7.50E-06	0	1	1	1	3	
.
0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	7	2.22E-09	1	1	1	0	3	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.0253	1	1	1	1	4	
0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	2	0.0017	1	1	1	1	4	
.
0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	3	0.00011	1	1	1	1	4	
.
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	8.67E-19	1	1	1	1	4	
															Sum	1						

Table 11. Probabilities by brute force.

Supplement (3): strategy to decipher the results of two, three and four positives

When test result is 0011, for example, it implies specimen 3 is positive as would be in the hypothetical scenario. It is true that specimen 3 must be positive if both specimens 1 and 2 are negative. Therefore, we can pool specimens 1 and 2 together for testing. If negative, then specimen 3 is positive. If positive, then up to three additional tests are needed.

The probability to get two positive groups but only one positive specimen is 15.19%. (See table 12.) Therefore, 15.19% of the time, five group tests would return an answer. And 3.11% of the time, total of eight tests are needed. On average, the number of tests is: $15.19\% \times 5 + 3.11\% \times 8 = 1.0083$

Similarly, test result 0111 implies specimen 7 is positive if specimens 1 to 6 are negative. The probability to get three positive groups but only one positive specimen is 10.13%. (See table 12.) Therefore, 10.13% of the time, five group tests would return an answer. And 9.65% of the time, total of 12 tests are needed. On average, the number of tests is: $10.13\% \times 5 + 9.65\% \times 12 = 1.6645$

For the same reason, test result 1111 implies specimen 15 is positive if specimens 1 to 14 are negative. The probability to get three positive groups but only one positive specimen is 2.53%. (See table 12.) Therefore, 2.53% of the time, five group tests would return an answer. And 11.28% of the time, total of 20 tests are needed. On average, the number of tests is: $2.53\% \times 5 + 11.28\% \times 20 = 2.3825$

Overall, total number of tests:

$$(0.3798 + 0.1013) \times 4 + 1.0083 + 1.6645 + 2.3825 = 6.9797$$

Total number of time cycles:

$$(0.3798 + 0.1013) \times 1 + (0.1519 + 0.1013 + 0.0253) \times 2 + (0.0311 + 0.0965 + 0.1128) \times 3 = 1.7593$$

Number of positives	Probability	Number of positive specimens				
		0	1	2	3	4
Zero	37.98%	37.98%	0	0	0	0
One	10.13%		10.13%	0	0	0
Two	18.30%		15.19%	3.04%	0.07%	0
Three	19.78%		10.13%	8.10%	1.44%	...
Four	13.81%		2.53%	6.58%	3.61%	...
Total	100%	37.98%	37.98%	17.72%	5.12%	...

Table 12. Breakdown of probabilities.