Examining Two-stage Group Testing

Bo Li and David Saad

Non-linearity and Complexity Research Group, Aston University, Birmingham, B4 7ET, United Kingdom

June 9, 2020

The two-stage group testing method is examined and its application in COVID-19 is discussed. The expected false positive rate and false negative rate are calculated for given values of sensitivity, specificity, prevalence and group-testing parameters. To reduced false negative rate, we suggest to test the pooled sample multiple times at the first stage.

1 Procedures and Efficacy of Two-stage Group Testing

The task is to test N individuals to determine whether each individual is infected by some disease or not by using the two-stage group testing strategy. An overview of group testing can be found in [1]. Denote p as the infection probability, which is the prevalence parameter of the epidemics. Denote ρ, q as the sensitivity and specificity of the test, respectively, which are parameters of the test itself.[1] The procedures of two-stage group test is described in Algorithm 1.

Such method has been adopted in Israel [2], Nebraska and Germany [3]. There are also studies which discuss the application of different group testing algorithms in COVID-19 [4, 5, 6].

The motivation of testing each pool for m times in Step 7 of Algorithm 1 is due to the concern of high false negative rate of the RT-PCR test for COVID-19 [7, 8]. If a pool contains several infected specimens but was falsely tested as negative, then all the infected individuals in this pool will be misclassified, which effectively increases the false negative rate even further. More precisely, we found (see the theory section below) that introducing the two-stage group testing improves the performance of testing positives (positive predictive value PPV becomes higher and false positive rate FPR becomes

[1] The sensitivity ρ indicates the probability of obtaining a positive result when testing on an infected individual, while the specificity q indicates the probability of obtaining a negative result when testing on a healthy individual.

Algorithm 1 Procedures of two-stage group test.

Inputs: N specimens, sensitivity ρ, specificity q, estimated prevalence parameter p

1: On the basis of the developed theory, determine the pool size k, number of pools N/k, and number of tests m performed on each pool.
2: Partition the N specimens into N/k non-overlapping pools, where each pool has k specimens.
3: Divide each specimen into $m+1$ fractions.
4: **for** $i = 1$ to N/K **do**
5: In the i-th pool, take a specimen fraction of each individual and mix them into one big sample.
6: Test the big sample of the pool and record the test result.
7: Repeat Step 5 and Step 6 for m times.
8: **if** all the m tests of the pool are negative **then**
9: Mark all k individuals of the i-th pool as negative.
10: **else**
11: Test all k specimens of the i-th pool individually and record the results.
12: **end if**
13: **end for**

Outputs: test results of N specimens.

lower), while it degrades the performance of testing negatives (negative predictive value NPV becomes lower and false negative rate FNR becomes higher).

To determine the parameters in Step 1 of Algorithm 1, one way is to minimize the expected number of tests T, subject to some accuracy constraints. To be more specific, we consider the case $\rho = 0.9, q = 0.999$ (high specificity but moderate sensitivity), which is relevant in testing of COVID-19. For individual test, the false negative rate $FNR = 1-\rho = 10\%$. We found in theory (Eq. (24) below) that for two-stage group test, expected FNR only depends on ρ and m. For group test with $m = 1$, $FNR = 1 - \rho^2 = 19\%$, which is much higher than the case of individual test. However, if $m = 2$ is used, then expected $FNR = 10.9\%$; if $m = 3$ is used, expected $FNR = 10.09\%$. Therefore, testing a pool multiple times can reduce the FNR to similar level of individual test. Generally speaking, a higher m is needed for a lower sensitivity ρ. For $\rho = 0.9$, we suggest to choose $m = 2$ or $m = 3$, and minimize number of tests T in Eq. (20) with respect to k. An example is shown in Table. 1. *It is shown that a significant number of tests can be saved while maintaining a similar level of statistical performance.*

Notice that in developing the theory, we assume the sensitivity ρ and specificity q are constant irrespective of sample dilution, which needs to be examined [9, 10, 11, 12].

2 Theory and Determination of Parameters

In this section, we outline the theory of two-stage group test, which can be used to determine the parameters needed for implementation. Note that only average behaviors

p	m	k	T/N	PPV	NPV	FPR	FNR
0.01		1	100%	0.9009	0.99899	0.001	0.1
0.01	2	15	27.3%	0.9856	0.99890	0.0001317	0.109
0.01	3	19	33.4%	0.9819	0.99898	0.0001678	0.1009
0.02		1	100%	0.9484	0.99796	0.001	0.1
0.02	2	11	38.1%	0.9901	0.99778	0.0001827	0.109
0.02	3	14	46.3%	0.9875	0.99794	0.0002331	0.1009

Table 1: Optimal parameters of group testing, compared to individual testing ($k = 1$). Parameters are $\rho = 0.9, q = 0.999$. A significant number of tests can be saved while maintaining a similar level of statistical performance. For example, when $p = 0.01$, using group-testing with parameters $m = 2, k = 15$ only requires 27.3% testing resources of individual testing, while keeping a similar level of FPR and FNR.

are studied, while fluctuations are not considered. Some quantities of interests are

$$T = \mathbb{E}(\text{number of tests}), \tag{1}$$

$$\begin{aligned} PPV = \text{positive predictive value} &= \frac{\mathbb{E}(\text{number of true test positives})}{\mathbb{E}(\text{number of test positives})} \\ &= \frac{\mathbb{E}(\text{number of true test positives})}{\mathbb{E}(\text{number of true test positives}) + \mathbb{E}(\text{number of false positives})}, \end{aligned} \tag{2}$$

$$\begin{aligned} NPV = \text{negative predictive value} &= \frac{\mathbb{E}(\text{number of true test negatives})}{\mathbb{E}(\text{number of test negatives})} \\ &= \frac{\mathbb{E}(\text{number of true test negatives})}{\mathbb{E}(\text{number of true test negatives}) + \mathbb{E}(\text{number of false negatives})}, \end{aligned} \tag{3}$$

$$FPR = \text{false positive rate} = \frac{\mathbb{E}(\text{number of false positives})}{\mathbb{E}(\text{number of healthy individuals})}, \tag{4}$$

$$FNR = \text{false negative rate} = \frac{\mathbb{E}(\text{number of false negatives})}{\mathbb{E}(\text{number of infected individuals})}. \tag{5}$$

If the infection probability (prevalence parameter) is p, we have

$$\mathbb{E}(\text{number of infected individuals}) = Np, \tag{6}$$

$$\mathbb{E}(\text{number of healthy individuals}) = N(1-p). \tag{7}$$

2.1 Theory of Individual Test

For conventional individual test method (i.e., $k = 1$), each individual is tested once. The quantities of interests are computed as

$$T = N, \tag{8}$$

$$PPV = \frac{p\rho}{p\rho + (1-p)(1-q)}, \tag{9}$$

$$NPV = \frac{(1-p)q}{(1-p)q + p(1-\rho)}, \tag{10}$$

$$FPR = 1 - q, \tag{11}$$

$$FNR = 1 - \rho. \tag{12}$$

2.2 Theory of Group Test

For two-stage group testing (i.e., $k > 1$) as described, some relevant quantities are computed as

$$\mathbb{E}(\text{number of infected groups}) = \frac{N}{k}\big[1 - (1-p)^k\big], \tag{13}$$

$$\mathbb{E}(\text{number of healthy groups}) = \frac{N}{k}(1-p)^k, \tag{14}$$

$$\begin{aligned}&\mathbb{E}(\text{number of infected individuals inside an infected group})\\ =&\frac{\mathbb{E}(\text{number of infected individuals})}{\mathbb{E}(\text{number of infected groups})} = \frac{Np}{\frac{N}{k}\big[1-(1-p)^k\big]} = \frac{kp}{1-(1-p)^k},\end{aligned} \tag{15}$$

$$\begin{aligned}&\mathbb{E}(\text{number of true test positives})\\ =&\frac{N}{k}\big[1-(1-p)^k\big]\cdot\big[1-(1-\rho)^m\big]\cdot\frac{kp}{1-(1-p)^k}\rho\\ =&Np\cdot\big[1-(1-\rho)^m\big]\cdot\rho,\end{aligned} \tag{16}$$

$$\begin{aligned}&\mathbb{E}(\text{number of false positives})\\ =&\frac{N}{k}(1-p)^k\cdot(1-q^m)\cdot k(1-q)\\ &+\frac{N}{k}\big[1-(1-p)^k\big]\cdot\big[1-(1-\rho)^m\big]\cdot\left[k-\frac{kp}{1-(1-p)^k}\right](1-q)\\ =&N(1-p)^k(1-q^m)(1-q)+N\big[1-p-(1-p)^k\big]\cdot\big[1-(1-\rho)^m\big](1-q),\end{aligned} \tag{17}$$

$$
\begin{aligned}
&\mathbb{E}(\text{number of true test negatives}) \\
=&\frac{N}{k}(1-p)^k \cdot q^m \cdot k + \frac{N}{k}(1-p)^k \cdot (1-q^m) \cdot kq \\
&+ \frac{N}{k}\big[1-(1-p)^k\big] \cdot (1-\rho)^m \cdot \left[k - \frac{kp}{1-(1-p)^k}\right] \\
&+ \frac{N}{k}\big[1-(1-p)^k\big] \cdot \big[1-(1-\rho)^m\big] \cdot \left[k - \frac{kp}{1-(1-p)^k}\right] q \\
=&N(1-p)^k\big[q^m + (1-q^m)q\big] \\
&+ N\big[1-p-(1-p)^k\big] \cdot \Big[(1-\rho)^m + \big[1-(1-\rho)^m\big] \cdot q\Big],
\end{aligned} \tag{18}
$$

$$
\begin{aligned}
&\mathbb{E}(\text{number of false negatives}) \\
=&\frac{N}{k}\big[1-(1-p)^k\big] \cdot (1-\rho)^m \cdot \frac{kp}{1-(1-p)^k} \\
&+ \frac{N}{k}\big[1-(1-p)^k\big] \cdot \big[1-(1-\rho)^m\big] \cdot \frac{kp}{1-(1-p)^k} \cdot (1-\rho) \\
=&Np(1-\rho)^m + Np\big[1-(1-\rho)^m\big](1-\rho),
\end{aligned} \tag{19}
$$

based on which the quantities of interests are computed as

$$
T = \frac{N}{k} \cdot m + \left[\frac{N}{k}\big[1-(1-p)^k\big] \cdot \big[1-(1-\rho)^m\big] + \frac{N}{k}(1-p)^k \cdot (1-q^m)\right] k, \tag{20}
$$

$$
PPV = \frac{\mathbb{E}(\text{number of true test positives})}{\mathbb{E}(\text{number of true test positives}) + \mathbb{E}(\text{number of false positives})}, \tag{21}
$$

$$
NPV = \frac{\mathbb{E}(\text{number of true test negatives})}{\mathbb{E}(\text{number of true test negatives}) + \mathbb{E}(\text{number of false negatives})}, \tag{22}
$$

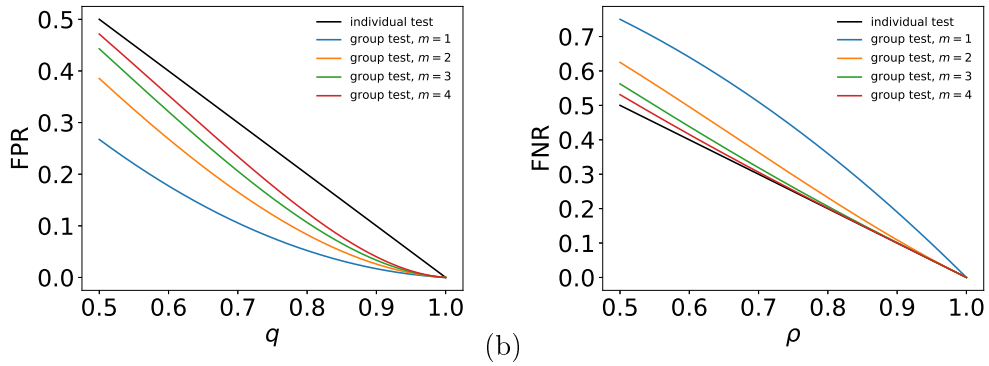
$$
FPR = \Big\{(1-p)^{k-1}(1-q^m) + \big[1-(1-p)^{k-1}\big] \cdot \big[1-(1-\rho)^m\big]\Big\} \cdot (1-q), \tag{23}
$$

$$
FNR = \Big\{(1-\rho)^{m-1} + \big[1-(1-\rho)^m\big]\Big\} \cdot (1-\rho). \tag{24}
$$

The group-testing parameters m, k can be determined based on these quantities of interests, primarily by minimizing T while controlling FPR and FNR, as was done in Table 1.

2.3 Some Observations

In Fig. 1, we demonstrate the effect of group testing on FPR and FNR. Comparing to individual testing, group testing reduces FPR, but increases FNR. However, FNR can be controlled with a relatively small m.



(a) (b)

Figure 1: (a) FPR vs q. Parameters are $p = 0.01, \rho = 0.9, k = 10$. Comparing to individual testing, group testing reduces FPR. FPR increases with m. (b) FNR vs ρ. Parameters other than ρ and m are irrelevant. Comparing to individual testing, group testing increases FNR. However, FNR can be controlled with a relatively small m.

In Fig. 2, we contrast the theory to simulation; it demonstrates a good match of average behaviors between the two approaches, while fluctuations are large. On average, group testing improves PPV but degrades NPV slightly comparing to individual testing; it also reduces FPR but increases FNR.

Acknowledgments

We acknowledge support from European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 835913.

References

[1] Matthew Aldridge, Oliver Johnson, and Jonathan Scarlett. Group testing: An information theory perspective. *Foundations and Trends in Communications and Information Theory*, 15(3-4):196–392, 2019.

[2] https://www.technion.ac.il/en/2020/03/pooling-method-for-accelerated-testing-of-covid-19/, 2020.

[3] https://www.nytimes.com/2020/05/07/opinion/coronavirus-group-testing.html, 2020.

[4] Christian Gollier and Olivier Gossner. Group testing against covid-19. http://www.lse.ac.uk/school-of-public-policy/assets/Documents/Social-Sciences-Response-to-Covid/Professor-Olivier-Gossner-One.pdf, 2020.

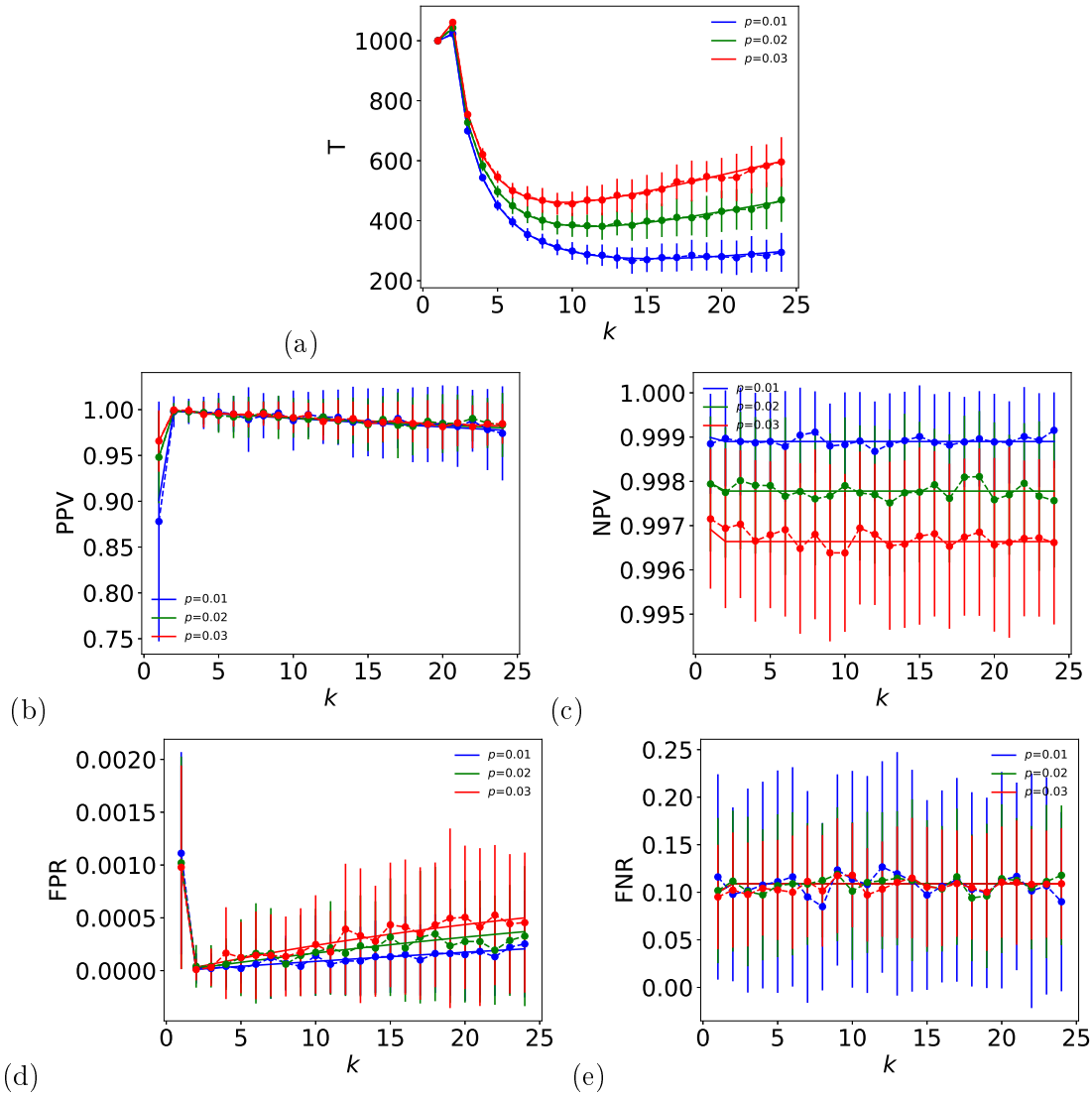


Figure 2: Solid line: theory; dash line with error bar: simulation (each data point is averaged over 100 realizations). Parameters are $N = 1000, \rho = 0.9, q = 0.999, m = 2$. (a) T vs k. Group testing reduces the expected number of tests required when $k > 2$. (b) PPV vs k. (c) NPV vs k. (d) FPR vs k. (e) FNR vs k.

[5] Lakshmi N. Theagarajan. Group testing for covid-19: How to stop worrying and test more. arXiv:2004.06306, 2020.

[6] N. Bottman, Y. Cooper, and F. Janda. A comparison of group testing architectures for covid-19 testing. arXiv:2005.03051, 2020.

[7] https://www.sciencedaily.com/releases/2020/05/200526173832.htm, 2020.

[8] Jessica Watson, Penny F Whiting, and John E Brush. Interpreting a covid-19 test result. *BMJ*, 369, 2020.

[9] Idan Yelin, Noga Aharony, Einat Shaer Tamar, Amir Argoetti, Esther Messer, Dina Berenbaum, Einat Shafran, Areen Kuzli, Nagham Gandali, Omer Shkedi, Tamar Hashimshony, Yael Mandel-Gutfreund, Michael Halberthal, Yuval Geffen, Moran Szwarcwort-Cohen, and Roy Kishony. Evaluation of COVID-19 RT-qPCR test in multi-sample pools. *Clinical Infectious Diseases*, 2020.

[10] Baha Abdalhamid, Christopher R Bilder, Emily L McCutchen, Steven H Hinrichs, Scott A Koepsell, and Peter C Iwen. Assessment of Specimen Pooling to Conserve SARS CoV-2 Testing Resources. *American Journal of Clinical Pathology*, 153(6):715–718, 04 2020.

[11] Claudio M. Verdun, Tim Fuchs, Pavol Harar, Dennis Elbrächter, David S. Fischer, Julius Berner, Philipp Grohs, Fabian J. Theis, and Felix Krahmer. Group testing for SARS-CoV-2 allows for up to 10-fold efficiency increase across realistic scenarios and testing strategies. May 2020.

[12] Alexandra Martin, Alexandre Storto, Barbara Andre, Allison Mallory, Remi Dangla, Benoit Visseaux, and Olivier Gossner. High-sensitivity covid-19 group testing by digital pcr. arXiv:2006.02908, 2020.