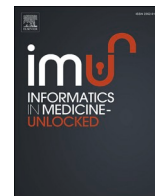




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Modeling the optimization of COVID-19 pooled testing: How many samples can be included in a single test?

Lu Liu^{*}

School of Economics, Southwestern University of Finance and Economics, China

ARTICLE INFO

Keywords:

Pooled testing
COVID-19
Resource cost function
Time function
Machine learning

ABSTRACT

Objectives: This study tries to answer the crucial question of how many biological samples can be optimally included in a single test for COVID-19 pooled testing.

Methods: It builds a novel theoretical model which links the local population to be tested in a region, the number of biological samples included in a single test, the “attitude” toward resource cost saving and time taken in a single test, as well as the corresponding resource cost function and time function, together. The numerical simulation results are then used to formulate the resource cost function as well as the time function. Finally, a loss function to be minimized is constructed and the optimal number of samples included is calculated.

Results: In a numerical example, we consider a region of 1 million population which needs to be tested for the infection of COVID-19. The solution calculates the optimal number of biological samples included in a single test as 4.254 when the time taken is given the weight of 50% under the infection probability of 10%. Other combinations of numerical results are also presented.

Conclusions: As we can see in our simulation results, given the infection probability at 10%, setting the number of biological samples included in a single test (in the integer level) at [4,6] is reasonable for a wide range of the subjective attitude between time and resource costs. Therefore, in the current practice, 5-mixed samples would sound better than the commonly used 10-mixed samples.

1. Introduction

Along with the rapid development of the COVID-19 pandemic around the globe, the need for testing the virus of SARS-CoV-2 at a massive level increases dramatically. The detection of SARS-CoV-2 RNA by Polymerase Chain Reaction (PCR) technology in testing is a common and sensitive method of COVID-19 diagnosis [1,2], even for reinfection [3]. In fact, whether such diagnostic tests are rapid and accurate enough or not can be the key to containing the pandemic [4,5].

After the outbreak in early 2020, to implement the full scale of work resumption in Wuhan, China, from May 15 to May 24, 2020, 9 million biological samples were collected, where >6.5 million were laboratory-tested [6,7]. In mid-August 2020, as school reopening is imminent, a massive level of COVID-19 tests with fast results was expected in the US [8]. On October 11, 2020, a small wave of outbreaks took place in Qingdao, China [9]. As a rapid response, the massive level population was quickly tested for COVID-19. A similar situation emerged in Chengdu, Shijiazhuang, Shenyang, and some other cities in China at the end of 2020. Mass testing is proven to be an efficient way to act as the

urgent response to disease control [10–14], especially at the screening stage [15–18], where frequency and turnaround time are more important than the test sensitivity [19]. In practice, this would reduce the shock to the economy due to the COVID-19 pandemic to a large extent [20,21].

In addition, the widespread COVID-19 testing in both clinical and non-clinical settings is also paid much attention [22–25]. Unfortunately, the failure to test the possible COVID-19 cases early enough would cause tragic results to contain outbreaks [26]. In fact, the limited availability of testing COVID-19 may be one important reason for the disastrous outcome of the pandemic in many places [27].

In the meantime, more and more studies emerge to examine the practical performance and testing strategies for SARS-CoV-2 [28,29], where the optimal testing strategies and maths-based strategies have begun to be discussed [30,31]. Now, the pooled testing technique became more and more popular since July 2020 in both the US and many other regions [32]. As self-explained by its name, COVID-19 pooled testing mixes several biological samples in a “batch”, and then tests the pooled samples with a diagnostic test. This approach increases

^{*} 555 Liutai Avenue, Wenjiang District, Chengdu, Sichuan, 611130, China.

E-mail address: liulu@swufe.edu.cn.

<https://doi.org/10.1016/j.imu.2022.101037>

Received 18 June 2022; Received in revised form 26 July 2022; Accepted 1 August 2022

Available online 7 August 2022

2352-9148/© 2022 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the number of samples to be tested using the same (or similar) amount of resources, hence reducing the medical cost and saving the time for a single test.

The advantage of pooled testing is obvious in the sense that we can save substantial time by examining a single test for mixed biological samples from multiple individuals. In addition, it can be quite cost-effective as well since fewer quantities of scarce resources, such as test kits, other testing supplies, and medical workers, are used. In fact, cost-effectiveness plays a central role in the strategies of SARS-CoV-2 testing from both medical and management perspectives [33,34]. Indeed, “fast and cheap” are essential to obtain and manage the tests [35–37]. Most importantly, it can boost the test capacity of a region dozens of times. For a simple example, if 10 biological samples are mixed for a single test, then theoretically the local test capacity for identifying the COVID-19 cases would be increased by about 10 times.

In the fourth quarter of 2020, the global daily increase of confirmed COVID-19 cases keeps at the one-million level. In many places all over the world, million-scale biological samples need to be tested every day. Although the multi-sample pools have already been academically studied [38–41], to our knowledge, the optimization of COVID-19 pooled testing is still not paid enough attention. In practice, it appears that the “rule of the thumb” prevails. Sometimes, 5 biological samples are mixed, while sometimes 10 samples or even more are combined. Then, how many samples can be optimally included in a single test? Easy as it appears to be, however, the answer to such a simple question is yet unknown. This study, therefore, tries to answer this crucial question and contribute to both the literature and practice. This paper attempted to obtain the optimal number of samples to be included in a single test for COVID-19 pooled testing, to balance the trade-offs between first-attempt costs and re-examine costs. Especially from a practical perspective, it studies an important problem to deal with the COVID-19.

2. Methods

2.1. The basic model

Let us start from a simple setting. Considering that a local population to be tested in a region is N_{Total} , and the number of biological samples included in a single test is G_{Test} , hence the number of tests can be denoted as $\frac{N_{Total}}{G_{Test}}$. In addition, we consider both the physical or monetary resource cost and the time taken in each test. Therefore, the optimization of COVID-19 pooled testing requires a tradeoff between the two aspects. Consequently, we introduce a parameter ζ to show the “attitude” toward resource cost saving. The corresponding “attitude” toward time-saving is thus $1-\zeta$. Moreover, we consider that the resource cost in monetary value for a single test is $C_{Test}(G_{Test})$, and the time taken in a single test is $T_{Test}(G_{Test})$. As we can see, we treat both C_{Test} and T_{Test} as functions of G_{Test} . This assumption is for a more general purpose, which can be simplified if both C_{Test} and T_{Test} are constants. However, we do see the nonlinear relationship in reality. Therefore, such a non-linear relationship may be more attractive.

Now, the total resource cost in the monetary value of the COVID-19 test in a region, C_{Total} , can be expressed as follows.

$$C_{Total} = \frac{N_{Total}}{G_{Test}} C_{Test}(G_{Test}) \quad (1)$$

The corresponding total time taken, T_{Total} , can be expressed as follows as well.

$$T_{Total} = \frac{N_{Total}}{G_{Test}} T_{Test}(G_{Test}) \quad (2)$$

Thus, we form a loss function to be minimized.

$$Loss = \zeta C_{Total} + (1-\zeta) T_{Total} \quad (3)$$

Substituting Eq. (1) and Eq. (2) into Eq. (3), and after simplification,

we have:

$$\min_{G_{Test}} Loss = \frac{N_{Total}}{G_{Test}} (\zeta C_{Test}(G_{Test}) + (1-\zeta) T_{Test}(G_{Test})) \quad (4)$$

Although the construction of the above loss function appears to be subjective, it is a reasonable representation of the trade-off between resource cost and time cost in COVID-19 testing. Solving the first-order condition from the above expression, we can obtain the optimal value of the number of samples in a single test below.

$$G_{Test}^* = \frac{\zeta C_{Test}(G_{Test}) + (1-\zeta) T_{Test}(G_{Test})}{\zeta \frac{\partial C_{Test}(G_{Test})}{\partial G_{Test}} + (1-\zeta) \frac{\partial T_{Test}(G_{Test})}{\partial G_{Test}}} \quad (5)$$

2.2. The extended model

Here, we introduce an extended model, which fits reality better. According to the standard protocol of the pooled testing of COVID-19, the positive or negative outcome of the pooled test would result in different operational treatments. Therefore, we need to consider the outcome of a single test, which would be “Yes” (i.e., positive) or “No” (i.e., negative). If “No”, then this test is passed. If “Yes”, then everybody in this group of samples would be tested again individually. In addition, we need to consider the “background” of the test, which is the infection intensity or severity of COVID-19 in a region. We thus introduce a probability of local COVID-19 infection, $P_{Infection}$. Now, the probability of the positive result in a single test, i.e., P_{Yes} , can be expressed as follows.

$$P_{Yes} = \frac{G_{Test}}{N_{Total}} P_{Infection} \quad (6)$$

This is a multiplication of probabilities. First, there is a “background” infection rate in the local population. Second, it is the ratio of the sampled group to be positively tested. Therefore, when we link them together, we obtain Eq. (6). Of course, such notation relies on the assumption of uniform distribution of the COVID-19 infection in a target group of the population. Then, the corresponding probability of the negative result in a single test, P_{No} , becomes $1-P_{Yes}$.

Notably, when the result in a single test is positive and everyone in the group needs to be re-examined, both the resource cost and time required in that group would increase dramatically. We hence denote the resource cost of re-examining a group to be $G_{Test} C_{Test}(G_{Test} = 1)$, where “($G_{Test}=1$)” means that the function “ $C_{Test}(G_{Test})$ ” is evaluated at the particular value of $G_{Test}=1$. Please consider “ $C_{Test}(\bullet)$ ” as a generalized function, then it would be easier to understand here. The corresponding time associated with the re-examination would be $G_{Test} T_{Test}(G_{Test} = 1)$ as well. As we can see, now both C_{Test} and T_{Test} are not functions of G_{Test} anymore as G_{Test} takes the specific numerical value of 1.

Finally, we can set up the mathematical expectation of the loss function as below.

$$E[Loss] = P_{Yes} \frac{N_{Total}}{G_{Test}} (\zeta G_{Test} C_{Test}(G_{Test} = 1) + (1-\zeta) G_{Test} T_{Test}(G_{Test} = 1)) + (1-P_{Yes}) \frac{N_{Total}}{G_{Test}} (\zeta C_{Test}(G_{Test}) + (1-\zeta) T_{Test}(G_{Test})) \quad (7)$$

Simplifying the above equation along with Eq. (6), we have:

$$E[Loss] = P_{Infection} (\zeta G_{Test} C_{Test}(G_{Test} = 1) + (1-\zeta) G_{Test} T_{Test}(G_{Test} = 1)) + (\frac{N_{Total}}{G_{Test}} - P_{Infection}) (\zeta C_{Test}(G_{Test}) + (1-\zeta) T_{Test}(G_{Test})) \quad (8)$$

In order to find the optimal solution, now we differentiate the above equation with respect to G_{Test} and we obtain the following expression.

$$\frac{\partial E[Loss]}{\partial G_{Test}} = P_{Infection} (\zeta C_{Test}(G_{Test} = 1) + (1-\zeta) T_{Test}(G_{Test} = 1)) - \frac{N_{Total}}{G_{Test}^2} (\zeta C_{Test}(G_{Test}) + (1-\zeta) T_{Test}(G_{Test})) + (\frac{N_{Total}}{G_{Test}} - P_{Infection}) (\zeta \frac{\partial C_{Test}(G_{Test})}{\partial G_{Test}} + (1-\zeta) \frac{\partial T_{Test}(G_{Test})}{\partial G_{Test}}) = 0 \quad (9)$$

Multiplying the term G_{Test}^2 on both sides of the above equation and after rearrangement, we have:

$$P_{Infection}(\zeta(C_{Test}(G_{Test} = 1) - \frac{\partial C_{Test}(G_{Test})}{\partial G_{Test}}) + (1 - \zeta)(T_{Test}(G_{Test} = 1) - \frac{\partial T_{Test}(G_{Test})}{\partial G_{Test}}))G_{Test}^2 + N_{Total}(\zeta \frac{\partial C_{Test}(G_{Test})}{\partial G_{Test}} + (1 - \zeta) \frac{\partial T_{Test}(G_{Test})}{\partial G_{Test}})G_{Test} - N_{Total}(\zeta C_{Test}(G_{Test}) + (1 - \zeta)T_{Test}(G_{Test})) = 0 \tag{10}$$

As shown above, the method used here follows the idea of artificial intelligence and especially machine learning, which starts from a loss function as Eq. (7) has shown. This strand of methods has already emerged in the study of COVID-19 pandemic-related issues [42–44]. Now, for the more general situation, without any form of simplification, we assume the following functional form.

$$C_{Test} = a + bG_{Test}^2 \tag{11}$$

We therefore have:

$$C_{Test}(G_{Test} = 1) = a + b \tag{12}$$

$$\frac{\partial C_{Test}(G_{Test})}{\partial G_{Test}} = 2bG_{Test} \tag{13}$$

Similarly, we assume:

$$T_{Test} = d + eG_{Test}^2 \tag{14}$$

And we have:

$$T_{Test}(G_{Test} = 1) = d + e \tag{15}$$

$$\frac{\partial T_{Test}(G_{Test})}{\partial G_{Test}} = 2eG_{Test} \tag{16}$$

Substituting the above expressions into Eq. (10), we have:

$$P_{Infection}(\zeta(a + b - 2bG_{Test}) + (1 - \zeta)(d + e - 2eG_{Test}))G_{Test}^2 + N_{Total}(\zeta 2bG_{Test} + (1 - \zeta)2eG_{Test})G_{Test} - N_{Total}(\zeta(a + bG_{Test}^2) + (1 - \zeta)(d + eG_{Test}^2)) = 0 \tag{17}$$

After rearrangement, we finally have:

$$2P_{Infection}(\zeta b + (1 - \zeta)e)G_{Test}^3 - (P_{Infection}(\zeta(a + b) + (1 - \zeta)(d + e)) + N_{Total}(\zeta b + (1 - \zeta)e))G_{Test}^2 + (N_{Total}\zeta a + N_{Total}(1 - \zeta)d) = 0 \tag{18}$$

As we know, the cubic equation is “notoriously” difficult to find a closed-form solution. Unfortunately, we cannot find a closed expression for the solution of the above equation. However, we can still use numerical methods to find an answer.

3. Data

As shown previously, the functional form of the resource cost for a single test $C_{Test}(G_{Test})$, and the time taken $T_{Test}(G_{Test})$ would be crucial. However, the available information on such functional form is very limited. Therefore, we have to derive them from both real-world data and simulation-based techniques.

For the resource cost of the test, we use the price of the test to stand for its overall resource cost. The trouble is the number of valid

observations. Using the latest information on the prices of single or pooled tests in China, we construct a data set of only 21 observations (Table 1).

In early 2021, 80 yuan is charged for a single-sample test, 30 yuan per person is charged for the pooled test mixed with 5 samples, and 20 or 15 yuan per person is charged for the pooled test mixed with 10 samples [45]. In the latest practice, the price charged has been reduced substantially. This is why we can see a downward-sloping trend in the associated prices of the PCR test.

Regarding the time taken for the tests, from a description by the medical worker [46], a single biological sample requires 2–4 h to be handled and a mix of 500 samples requires 8–10 h. Therefore, we construct a small data set with 6 observations from such description (Table 2).

Table 1
Constructed data set for the resource cost of COVID-19 testing.

C_{Test} Unit: RMB Yuan	G_{Test} Unit: Number of samples
80	1
150	5
200	10
150	10
60	1
58	1
90	5
130	10
40	1
40	5
80	10
35	1
100	5
50	5
150	10
16	1
25	5
50	10
6	1
17.5	5
35	10

Table 2
Constructed data set for the time of COVID-19 testing.

T_{Test} Unit: Number of hours	G_{Test} Unit: Number of samples
2	1
3	1
4	1
8	500
9	500
10	500

Table 3
Empirical estimation results for resource cost function of COVID-19 testing, with dependent variable C_{Test} ($N = 21$).

	Model (3-1) OLS	Model (3-2) LGM MCMC Draws: 10,000	Model (3-3) LGM MCMC Draws: 1,000,000	Model (3-4) LGM MCMC Draws: 10,000,000	Model (3-5) LGM MCMC Draws: 20,000,000
Constant	45.178*** (3.205)	15.923* (1.834)	15.879* (1.845)	15.878* (1.846)	15.882* (1.847)
G_{Test}^2	0.696*** (2.938)	1.044*** (5.346)	1.043148*** (5.337)	1.043*** (5.342)	1.043*** (5.338)
Adjusted R ²	0.276				

Note: t statistics in parentheses. *** $p \leq 0.01$, ** $0.01 < p < 0.05$, * $0.05 < p < 0.1$.

Table 4
Empirical estimation results for time function of COVID-19 testing, with dependent variable T_{Test} ($N = 6$).

	Model (4-1) OLS	Model (4-2) LGM MCMC Draws: 10,000	Model (4-3) LGM MCMC Draws: 1,000,000	Model (4-4) LGM MCMC Draws: 10,000,000	Model (4-5) LGM MCMC Draws: 20,000,000
Constant	3.000*** (5.196)	2.996*** (6.715)	2.994*** (6.693)	2.994*** (6.701)	2.994*** (6.705)
G_{Test}^2	2.40E-05*** (7.348)	2.40E-05*** (9.562)	2.40E-05*** (9.486)	2.40E-05*** (9.486)	2.40E-05*** (9.486)
Adjusted R ²	0.914				

Note: t statistics in parentheses. *** $p \leq 0.01$, ** $0.01 < p < 0.05$, * $0.05 < p < 0.1$.

4. Results

As discussed above, the estimation of the resource cost function and time function is challenging. We, therefore, take a two-step approach for each of the functions. First, we use Ordinary Least Square (OLS) upon the small-sample data set. Second, using the OLS results as the initial values, we simulate the large-sample results by the Linear Gaussian Model (LGM) with Gibbs sampling to apply the Markov Chain Monte Carlo (MCMC). These numerical simulation results are then used to formulate the resource cost function as well as the time function. Therefore, from the MCMC results with 20 million iterations, we obtain the numerical values as: $a=15.882$, $b=1.043$, $d=2.994$, $e=2.40E-05$ (Table 3, Table 4). Substituting these parameter values back into Eq. (18), we can have the full information of the important functions.

Now we can make a numerical example of the result, i.e., the optimal number of biological samples included in a single test. We consider a region of 1 million population who needs to be tested for the infection of COVID-19. Thus, besides the above functional parameter values, we also set the following numerical values: $N_{Total}=1,000,000$, $P_{Infection}=0.1$, $\zeta=0.5$. Numerically solving Eq. (18), we have $G_{Test}^*=4.254024$. Please note that the number of samples included in a single test should be an integer. Therefore, the calculated result can be made practical to the closest integer.

5. Discussion

The previous model and numerical results show the promise of conducting the COVID-19 pooled testing optimally both in theory and practice. Although there are already some other studies on a similar topic, the novelty of this study is likely to be the trade-off between cost

Table 5
Numerical simulation of G_{Test}^* .

	Model (5-1) $N_{Total}=10,000$ $P_{Infection}=0.1$ $\zeta=0.5$	Model (5-2) $N_{Total}=10,000,000$ $P_{Infection}=0.1$ $\zeta=0.5$	Model (5-3) $N_{Total}=1,000,000$ $P_{Infection}=0.1$ $\zeta=0.1$	Model (5-4) $N_{Total}=1,000,000$ $P_{Infection}=0.1$ $\zeta=0.9$
G_{Test}^*	4.253798	4.253532	6.407177	3.942798

and time constraints, which is comprehensively examined in the loss function mentioned in the above models.

However, when the exogenous conditions change, such calculated results would change accordingly. In Table 5, we show the new results when N_{Total} and ζ change their numerical values. As we can see, when N_{Total} becomes smaller, the number of samples included in a single test would be slightly larger. But it does not really make a difference in the practice level. On the contrary, the smaller ζ leads to a significantly larger value in G_{Test}^* . This result confirms our intuition that when the time taken in the test is given more weight in the decision (i.e., smaller value in ζ), there must be more samples included in a single test.

However, the discussion of $P_{Infection}$ is much more complex. Although we do not have a closed-form solution for G_{Test}^* now, we can still use the *Implicit Function Theorem* to derive the partial derivative of G_{Test} with respect to $P_{Infection}$ as below.

$$\frac{\partial G_{Test}}{\partial P_{Infection}} = \frac{(\zeta(a+b-2bG_{Test}) + (1-\zeta)(d+e-2eG_{Test}))G_{Test}^2}{P_{Infection}(\zeta(-2b) + (1-\zeta)(-2e))G_{Test}^2} + P_{Infection}(\zeta(a+b-2bG_{Test}) + (1-\zeta)(d+e-2eG_{Test}))2G_{Test} + N_{Total}(\zeta 2b + (1-\zeta)2e)G_{Test} + N_{Total}(\zeta 2bG_{Test} + (1-\zeta)2eG_{Test}) - N_{Total}(\zeta 2bG_{Test} + (1-\zeta)2eG_{Test}) \tag{19}$$

Simplifying the above equation and rearranging terms, we have:

$$\frac{\partial G_{Test}}{\partial P_{Infection}} = \frac{(\zeta(a+b-2bG_{Test}) + (1-\zeta)(d+e-2eG_{Test}))G_{Test}}{2P_{Infection}(\zeta b + (1-\zeta)e)G_{Test}} - 2P_{Infection}(\zeta(a+b-2bG_{Test}) + (1-\zeta)(d+e-2eG_{Test})) - 2N_{Total}(\zeta b + (1-\zeta)e) \tag{20}$$

We can easily locate a particular value where $\frac{\partial G_{Test}}{\partial P_{Infection}} = 0$ as follows.

$$G_{Test}^{***} = \frac{\zeta(a+b) + (1-\zeta)(d+e)}{2\zeta b + 2(1-\zeta)e} \tag{21}$$

Please note that such a particular value only means that $P_{Infection}$ does not influence G_{Test} . However, the expression of Eq. (20) also reveals the complex nature of such a relationship. It can be either positive or negative in a more general situation. Therefore, the sign of the above expression may be ambiguous.

6. Conclusion and future work

This study has proposed a theoretical model for the number of biological samples needed in a single test. It will help in resource and time minimization. As we can see in our simulation results, given the infection probability at 10%, setting the number of biological samples included in a single test (in the integer level) at [4,6] is reasonable for a wide range of the subjective attitude between time and resource costs. If we relax this criterion for a little bit, the range of [3,7] would also be OK

for the mixed biological samples from multiple individuals. Therefore, in the current practice, 5-mixed samples would sound better than the commonly used 10-mixed samples.

The method proposed in this study is of general purpose and it can be applied elsewhere other than in China. It can be used to analyze issues of other infectious diseases as well. One just needs to calibrate the functional parameters again with the local data, and the model would be still working. Although we may have made some novel contributions in this paper, this study still has some shortcomings.

On the one hand, concerns regarding the pooled testing arise in the sense that the virus concentration would decrease after being mixed with other samples. Indeed, sensitivity matters in the test [47,48]. Since samples are diluted, there is a higher likelihood to obtain false-negative results. Therefore, the chance of error in test results may be higher in mixed samples than that in a pure sample when other things are equal, which may result in higher infection risk as well as re-examine cost [49]. It is true in reality in China. In some less developed regions (or cities), 20 samples are pooled together in a single test, which is reported to be less accurate. In fact, some cases show that it may be one of the reasons for the delay in the detection of infected residents in the local area. Therefore, the expression of concern here is meaningful.

On the other hand, the data used in this study are still very limited. More information is needed to obtain the resource cost function and time function in the COVID-19 testing with higher accuracy. These issues among others are left for future research.

Funding

Supported by Guanghai Talent Project of Southwestern University of Finance and Economics.

Data availability statement

All data used in this study are publicly available.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I thank the editors and anonymous referees for their valuable comments. All remaining mistakes are my own.

References

- Gandhi RT, Lynch JB, del Rio C. Mild or moderate Covid-19. *N Engl J Med* 2020; 383:1757–66.
- Zhang Z, Bi Q, Fang S, et al. Insight into the practical performance of RT-PCR testing for SARS-CoV-2 using serological data: a cohort study. *Lancet Microbe* 2021;2(2):e79–87.
- Wyllie AL, Fournier J, Casanovas-Massana A, et al. Saliva or nasopharyngeal swab specimens for detection of SARS-CoV-2. *N Engl J Med* 2020;383:1283–6.
- Hansen CH, Michlmayr D, Gubbels SM, et al. Assessment of protection against reinfection with SARS-CoV-2 among 4 million PCR-tested individuals in Denmark in 2020: a population-level observational study. *Lancet* 2021;397(10280): 1204–12.
- Vandenberg O, Martiny D, Rochas O, et al. Considerations for diagnostic COVID-19 tests. *Nat Rev Microbiol* 2021;19:171–83.
- China News, <http://www.chinanews.com/gn/2020/05-25/9194183.shtml>.
- Liu L. Emerging study on the transmission of the novel coronavirus (COVID-19) from urban perspective: evidence from China. *Cities* 2020;103:102759.
- Rafiei Y, Mello MM. The missing piece — SARS-CoV-2 testing and school reopening. *N Engl J Med* 2020;383:e126.
- Xing YH, Ni W, Hu XW, et al. Rapid response to an outbreak in Qingdao, China. *N Engl J Med* 2020;383:e129.
- Studdert DM, Hall MA. Disease control, civil liberties, and mass testing — calibrating restrictions during the covid-19 pandemic. *N Engl J Med* 2020;383: 102–4.
- Skitttrall JP, Fortune MD, Jalal H, et al. Diagnostic tool or screening programme? Asymptomatic testing for SARS-CoV-2 needs clear goals and protocols. *Lancet Regional Health - Europe* 2021;1:100002.
- Strom M, Kristiansen MF, Christiansen DH, et al. Elimination of COVID-19 in the Faroe Islands: effectiveness of massive testing and intensive case and contact tracing. *Lancet Regional Health - Europe* 2021;1:100011.
- Rothe C. Multiple testing and the effect of NPIs on the spread of SARS-CoV-2. *Lancet Infect Dis* 2021;21(4):458. [https://doi.org/10.1016/S1473-3099\(20\)30935-X](https://doi.org/10.1016/S1473-3099(20)30935-X).
- Smyrklaki I, Ekman M, Lentini A, et al. Massive and rapid COVID-19 testing is feasible by extraction-free SARS-CoV-2 RT-PCR. *Nat Commun* 2020;11:4812.
- Xie X, Muruato AE, Zhang X, et al. A nanoluciferase SARS-CoV-2 for rapid neutralization testing and screening of anti-infective drugs for COVID-19. *Nat Commun* 2020;11:5214.
- Ramdas K, Darzi A, Jain S. ‘Test, re-test, re-test’: using inaccurate tests to greatly increase the accuracy of COVID-19 testing. *Nat Med* 2020;26:810–1.
- Chaibun T, Puenpa J, Ngamdee T, et al. Rapid electrochemical detection of coronavirus SARS-CoV-2. *Nat Commun* 2021;12:802.
- He X, Quan S, Xu M, et al. Generation of SARS-CoV-2 reporter replicon for high-throughput antiviral screening and testing. *Proc Natl Acad Sci USA* 2021;118(15): e2025866118.
- Larremore DB, Wilder B, Lester E, et al. Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. *Sci Adv* 2021;7(1):eabd5393.
- Wang Q, Liu L. Pandemic or panic? A firm-level study on the psychological and industrial impacts of COVID-19 on the Chinese stock market. *Financial Innov* 2022; 8(1):36. <https://doi.org/10.1186/s40854-022-00335-8>.
- Liu L, Huang J, Li H. Estimating the real shock to the economy from COVID-19: the example of electricity use in China. *Technol Econ Dev Econ* 2022;1–21. <https://doi.org/10.3846/tede.2022.17027>.
- Boehme C, Hannay E, Sampath R. SARS-CoV-2 testing for public health use: core principles and considerations for defined use settings. *Lancet Global Health* 2021;9 (3):e247–9.
- Fitzpatrick MC, Pandey A, Wells CR, et al. Buyer beware: inflated claims of sensitivity for rapid COVID-19 tests. *Lancet* 2021;397(10268):24–5.
- MacKay MJ, Hooker AC, Afshinnekoo E, et al. The COVID-19 XPRIZE and the need for scalable, fast, and widespread testing. *Nat Biotechnol* 2020;38:1021–4.
- Rubin R. COVID-19 testing moves out of the clinic and into the home. *JAMA* 2021; 326(14):1362–4.
- Schneider EC. Failing the test — the tragic data gap undermining the U.S. Pandemic response. *N Engl J Med* 2020;383:299–302.
- Tu YP, Jennings R, Cangelosi GA, et al. Swabs collected by patients or health care workers for SARS-CoV-2 testing. *N Engl J Med* 2020;383:494–6.
- Zhang Z, Bi Q, Fang SS, et al. Insight into the practical performance of RT-PCR testing for SARS-CoV-2 using serological data: a cohort study. *Lancet Microbe* 2021;2:e79–87.
- Quilty BJ, Clifford S, Hellewell J, et al. Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. *Lancet Public Health* 2021;6(3): e175–83.
- Wells CR, Townsend JP, Pandey A, et al. Optimal COVID-19 quarantine and testing strategies. *Nat Commun* 2021;12:356.
- Nature Briefing. 28 October 2020. *Nat* 2020; 586: 64.
- CNN Health. <https://edition.cnn.com/2020/07/06/health/coronavirus-pool-testin-g-wellness/index.html>.
- Du ZW, Pandey A, Bai Y, et al. Comparative cost-effectiveness of SARS-CoV-2 testing strategies in the USA: a modelling study. *Lancet Public Health* 2021;6(3): e184–91.
- Mina MJ, Peto TE, García-Fiñana M, et al. Clarifying the evidence on SARS-CoV-2 antigen rapid tests in public health responses to COVID-19. *Lancet* 2021;397 (10283):1425–7.
- Peeling R, Olliaro P. Rolling out COVID-19 antigen rapid diagnostic tests: the time is now. *Lancet Infect Dis* 2021;21(8):1052–3.
- Boum Y, Fai KN, Nikolay B, et al. Performance and operational feasibility of antigen and antibody rapid diagnostic tests for COVID-19 in symptomatic and asymptomatic patients in Cameroon: a clinical, prospective, diagnostic accuracy study. *Lancet Infect Dis* 2021;21(8):1089–96.
- Mina MJ, Andersen KG. COVID-19 testing: one size does not fit all. *Science* 2021; 371(6525):126–7.
- Yelin I, Aharoni N, Tamar ES, et al. Evaluation of COVID-19 RT-qPCR test in multi sample pools. *Clin Infect Dis* 2020;71(16):2073–8.
- Aleta A, Martín-Corral D, Pastore y Piontti A, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat Human Behav* 2020;4:964–71.
- Barak N, Ben-Ami R, Sido T, et al. Lessons from applied large-scale pooling of 133,816 SARS-CoV-2 RT-PCR tests. *Sci Transl Med* 2021;13(589):eabf2823.
- Shental N, Levy S, Wuvshet V, et al. Efficient high-throughput SARS-CoV-2 testing to detect asymptomatic carriers. *Sci Adv* 2020;6(37):eabc5961.
- Soltan AAS, Kouchaki S, Zhu TT, et al. Rapid triage for COVID-19 using routine clinical data for patients attending hospital: development and prospective validation of an artificial intelligence screening test. *Lancet Digital Health* 2021;3: e78–87.
- Afrash MR, Kazemi-Arpanahi H, Shanbehzadeh M, et al. Predicting hospital readmission risk in patients with COVID-19: a machine learning approach. *Inform Med Unlocked* 2022;30:100908.
- Alabbad DA, Almuhaideb AM, Alsunaidi SJ, et al. Machine learning model for predicting the length of stay in the intensive care unit for Covid-19 patients in the eastern province of Saudi Arabia. *Inform Med Unlocked* 2022;30:100937.

- [45] The paper, https://www.thepaper.cn/newsDetail_forward_11322518.
- [46] Heilongjiang Daily. <http://epaper.hljnews.cn/hljrb/20200510/470583.html>.
- [47] Mina MJ, Parker R, Larremore DB. Rethinking covid-19 test sensitivity — a strategy for containment. *N Engl J Med* 2020;383:e120.
- [48] Congrave-Wilson Z, Lee Y, Jumarang J, et al. Change in saliva RT-PCR sensitivity over the course of SARS-CoV-2 infection. *JAMA* 2021;326(11):1065–7.
- [49] Coffey KC, Diekema DJ, Morgan DJ. Interpreting SARS-CoV-2 test results. *JAMA* 2021;326(15):1528–9.