CrossMark

# A discussion on the minimum required number of tests in two common pooling test methods for SARS-CoV-2

**Author for correspondence:**
W. K. Chow,
E-mail: wan-ki.chow@polyu.edu.hk

## C. H. Cheng[1], C. L. Chow[1] and W. K. Chow[2] (iD)

[1]Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, China and
[2]Department of Building Services Engineering, The Hong Kong Polytechnic University, Hong Kong, China

## Abstract

Pooling of samples in detecting the presence of virus is an effective and efficient strategy in screening carriers in a large population with low infection rate, leading to reduction in cost and time. There are a number of pooling test methods, some being simple and others being complicated. In such pooling tests, the most important parameter to decide is the pool or group size, which can be optimised mathematically. Two pooling methods are relatively simple. The minimum numbers required in these two tests for a population with known infection rate are discussed and compared. Results are useful for identifying asymptomatic carriers in a short time and in implementing health codes systems.

**CAMBRIDGE**
UNIVERSITY PRESS

## Introduction

Owing to the recent COVID-19 pandemic, there is a need to test a huge number of people in order to identify the SARS-CoV-2 carriers because many of them are asymptomatic. Even though a number of vaccines have been approved for emergency use, mass screening test is still needed in the near future before COVID-19 is under control. Mass screening test serves to identify asymptomatic carriers and stop potential transmission chains as soon as possible. Such a strategy is especially important in regions where vaccines are short of supply due to various causes, including limited global production capacity, politics and lack of strong global solidarity, all constituting obstacles to equitability in vaccine distribution [1, 2]. In addition to vaccine supply, an in-depth analysis of the unequal distribution of tests, identified cases and fatality rates in 18 most COVID-19-affected countries was recently reported by Shams *et al.* [3]. Furthermore, the effects of the COVID-19 pandemic on daily life were summarised by Haleem *et al.* [4]. Even in regions where vaccine supply is sufficient, the vaccination rate could be low due to poor understanding and hence irrational worrying of the side effects of vaccination among the general public. This is exactly the case in Hong Kong where the COVID-19 full vaccination rate is below 20% as of 12 June 2021 [5]. Mass screening test exerts tremendous pressure on the government in terms of resources and time. One possible way out to more efficiently carry out mass testing [6–14] is to make use of group testing via pooling samples from a group of people, aiming at saving time, materials and test kits. This is especially important in developing countries with limited resources as pointed out before [12–14]. With more efficient use of available resources, a wider testing coverage can be achieved, which is very important in COVID-19 control. By analysing the data in 39 countries, Wei *et al.* [15] showed that testing coverage is an important means of early identification of carriers and hence better control of the disease.

Owing to the high infectivity and high mutation rate of SARS-CoV-2, rapid identification and isolation of individuals infected with the virus via testing would be an ongoing process [8]. However, normal daily life and economic activities were greatly affected [4]. To alleviate such adverse effects, a modified strategy in cutting off the transmission chain of infection is to lock down a building and require the residents in that building for compulsory virus testing when the number of confirmed cases in that building reaches a certain value, as practised in Hong Kong [16]. The threshold number of confirmed cases to elicit lockdown depends on the strain of the virus and on whether the infection source is known. This method is effective, but it usually takes more than 12 h to complete the screening test for buildings in Hong Kong, which are usually tall buildings with large number of residents. Thus it is of paramount importance if the test can be completed in a shorter time so as to restore normal livelihood and business as soon as possible. In this sense, even when resources are not a problem, pooling test is of great value because it saves waiting time. The method used in Hong Kong could be transferred to heavily populated urban areas in big cities.

The method of pooling test has been widely validated for SARS-CoV-2 [17–20] as listed in Pikovski and Bentele [21]. Pooling tests were employed in mass screening of COVID-19 in

different countries over the world, including the USA, India, Brazil, China, etc., as summarised in a recent review by Grobe *et al.* [22]. There are many methods for group testing [6–8]. A brief summary on four feasible methods has been recently reviewed [9]. Among these, the simplest method, which is also the pooling method commonly used, was discussed earlier [12–14]. It is interesting to note that pooling test was also studied using mathematical modelling and simulation [23]. Infection rate, test characteristics, population size and testing capacity were used as input parameters and the output fields were the test time and test number required, the number of cases identified and the number of false positives.

In this short communication, another pooling method was studied and compared with the previous one with respect to the minimum number of tests required.

## Two simple pooling test methods

We discussed two of the pooling methods, most attractive by virtue of their simplicity [9] in this paper, with focus on the minimum number of tests required for a given infection rate. These two methods were called Method 1 and Method 2 among the four methods by Mallapaty [9].

In Method 1, the population under test is divided into groups of equal numbers of samples (say $n$ in each group). The $n$ samples are pooled together and tested. If a group is tested positive, then the $n$ samples in that group are tested individually.

Taking an example with one positive case in $m = 27$ samples. By using a group size of $n = 9$, the first round requires three tests and the second round requires nine tests, making a total of 12 *vs.* 27 tests without group testing.

In Method 2, the first round is the same as in Method 1, but the positive group is divided into subgroups of size $k$. Referring to the example of 27 samples, the positive group is divided into three subgroups each containing three samples in the second round. The samples in the positive subgroup are tested individually in the third round of testing. Thus, the total number of tests in the three rounds is 9, further reducing the number of tests.

It should be pointed that in the example above, the number 27 is used to demonstrate the methods in simple numbers. In reality, one positive case in 27 samples is a serious scenario of population infection, meaning that 3.7% of the population is infected. The advantage of pooling test is more pronounced when the infection rate is low, as was shown in the following analysis.

The value of $m$, that is, the average number of samples that contains one positive case, is determined by the current population infection rate $\theta$, with $m = 1/\theta$. In Method 1, the total number of tests $L$ depends on the value of $n$. In Method 2, $L$ depends on $n$ and $k$. A problem of practical importance is the choice of $n$ (and also $k$ in Method 2) to make $L$ a minimum. While the solution for Method 1 has been reported by the same authors [13], with $n = m^{1/2} = (1/\theta)^{1/2}$, optimisation for Method 2 has not been reported. Thus, it is the aim of this short communication to show how the minimum number of tests required can be achieved via suitable choice of group size in Method 2, and to compare its efficiency with Method 1.

The two pooling test methods discussed above are summarised in the flowcharts in Figure 1.

## Optimisation

Let $P$ be the population under test, $\theta$ be the current population infection rate. Then most probably there will be one positive case in $m = 1/\theta$ people. Further, let $n$ be the group size in the first round and $k$ be the subgroup size in the second round. The total number of tests $L$ is given by

$$L = \frac{P}{n} + \frac{P}{m}\left(k + \frac{n}{k}\right),$$

First-order partial derivatives:

$$\frac{\partial L}{\partial k} = \frac{P}{m}\left(1 - \frac{n}{k^2}\right), \frac{\partial L}{\partial n} = -\frac{P}{n^2} + \frac{P}{km} = P\left(\frac{1}{mk} - \frac{1}{n^2}\right),$$

First-order condition for critical points:

$$\frac{\partial L}{\partial k} = \frac{\partial L}{\partial n} = 0,$$

$$\frac{P}{m}\left(1 - \frac{n}{k^2}\right) = 0, \tag{1}$$

$$P\left(\frac{1}{mk} - \frac{1}{n^2}\right) = 0. \tag{2}$$

Solving for (1), one has

$$n = k^2. \tag{3}$$

Putting (3) into (2) yields

$$n^2 = mk, \ k^4 - mk = 0.$$

Solving for $k$, one has

$$k = \sqrt[3]{m}. \tag{4}$$

Putting (4) into (3) yields

$$n = m^{2/3}. \tag{5}$$

The critical point is $k = \sqrt[3]{m}$, $n = m^{2/3}$.
Second-order partial derivatives:

$$\frac{\partial^2 L}{\partial k^2} = \frac{2Pn}{mk^3}, \ \frac{\partial^2 L}{\partial n^2} = \frac{2P}{n^3}, \ \frac{\partial^2 L}{\partial k \partial n} = \frac{\partial^2 L}{\partial n \partial k} = -\frac{P}{mk^2}. \tag{6}$$

Second-order condition for checking critical points: Hessian matrix $D$ is defined by:

$$D = \begin{bmatrix} \dfrac{\partial^2 L}{\partial k^2} & \dfrac{\partial^2 L}{\partial k \partial n} \\ \dfrac{\partial^2 L}{\partial n \partial k} & \dfrac{\partial^2 L}{\partial n^2} \end{bmatrix} = \begin{bmatrix} \dfrac{2Pn}{mk^3} & -\dfrac{P}{mk^2} \\ -\dfrac{P}{mk^2} & \dfrac{2P}{n^3} \end{bmatrix}. \tag{7}$$

At critical point $k = \sqrt[3]{m}$, $n = m^{\frac{2}{3}}$, $D$ is given by:

$$D = \begin{bmatrix} \dfrac{2P}{m^{4/3}} & -\dfrac{P}{m^{5/3}} \\ -\dfrac{P}{m^{5/3}} & \dfrac{2P}{m^2} \end{bmatrix}. \tag{8}$$
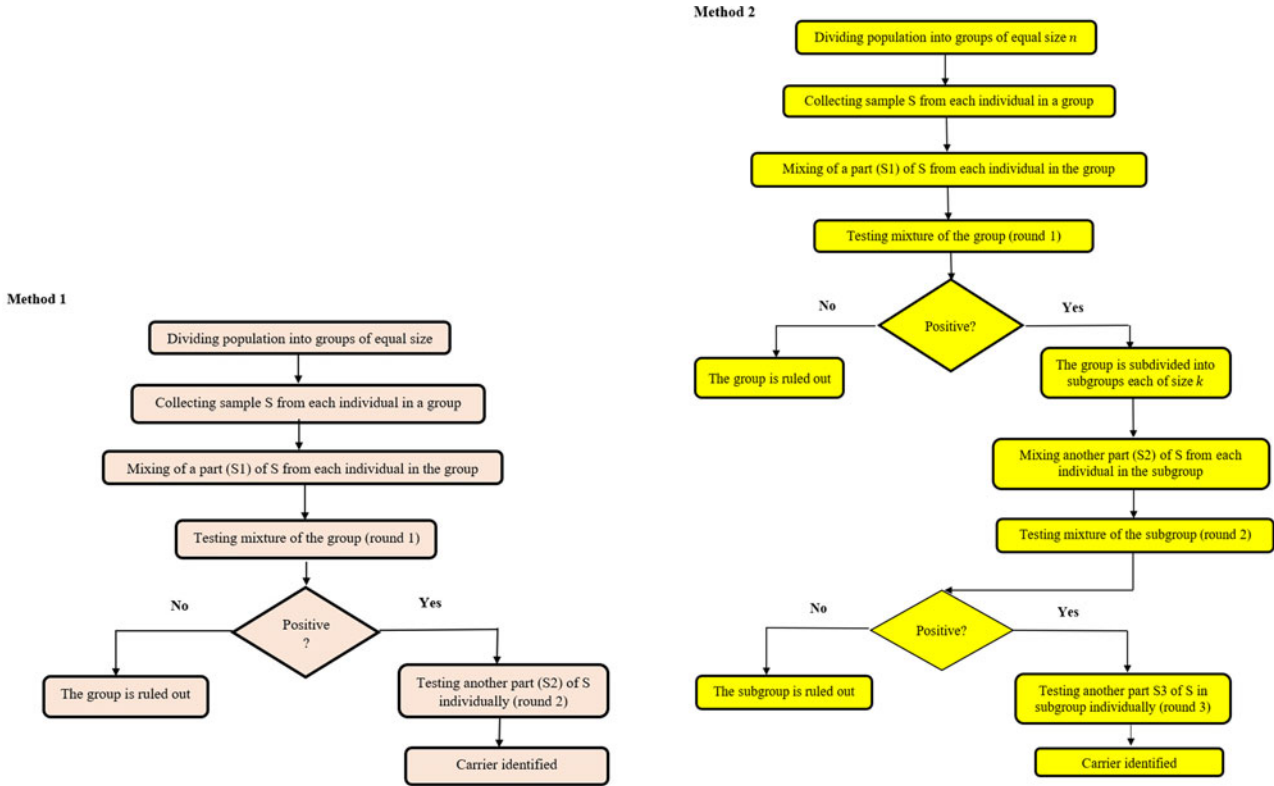
**Fig. 1.** Flowcharts summarizing the procedures in Method 1 and Method 2 in pooling test.

Sign-checking of $|D|$, $\partial^2 L/\partial k^2$ and $\partial^2 L/\partial n^2$ at critical point $k = \sqrt[3]{m}$, $n = m^{\frac{2}{3}}$:

$$|D| = \begin{vmatrix} \dfrac{2P}{m^{4/3}} & -\dfrac{P}{m^{5/3}} \\ -\dfrac{P}{m^{5/3}} & \dfrac{2P}{m^2} \end{vmatrix} = P \begin{vmatrix} \dfrac{2}{m^{4/3}} & -\dfrac{1}{m^{5/3}} \\ -\dfrac{1}{m^{5/3}} & \dfrac{2}{m^2} \end{vmatrix}$$

$$= P^2\left(\dfrac{4}{m^{10/3}} - \dfrac{1}{m^{10/3}}\right) = \dfrac{3P^2}{m^{10/3}} > 0, \qquad (9)$$

$$\dfrac{2P}{m^{4/3}} > 0, \quad \dfrac{2P}{m^2} > 0. \qquad (10)$$

Thus, the critical point $k = \sqrt[3]{m}$, $n = m^{\frac{2}{3}}$ is a local minimum point.

The minimum required number of tests ($L_{\min}$) in terms of $m$ at the critical point $k = \sqrt[3]{m}$, $n = m^{\frac{2}{3}}$ is given by

$$L_{\min} = \dfrac{P}{m^{2/3}} + \dfrac{P}{m}(2\sqrt[3]{m}) = \dfrac{3P}{m^{2/3}}. \qquad (11)$$

## Comparison of the two methods

A comparison of the two methods is given in Table 1 (for a population of $P$) at a population infection rate of $\theta$ and $m = 1/\theta$.

Numerical values are taken with reference to the observed number of infection cases. Taking a higher population infection

**Table 1.** Comparison of the number of tests using different methods

| Method | No. of tests $L$ | Min. No. of tests $L_{\min}$ |
|---|---|---|
| Without grouping | $L_o = P$ | $L_o = P$ |
| Method 1 (group size $= n$) | $L_1 = P\left(\frac{1}{n} + \frac{n}{m}\right)$ | $L_{1\min} = 2P/\sqrt{m}$, at $n = \sqrt{m}$ [12] |
| Method 2 (group size $= n$, subgroup size $= k$) | $L_2 = \frac{P}{n} + \frac{P}{m}\left(k + \frac{n}{k}\right)$ | $L_{2\min} = \frac{3P}{m^{2/3}}$, at $k = \sqrt[3]{m}$, $n = m^{\frac{2}{3}}$ |

rate of identifying 1 AC in a building instructed to have mandatory screening of 1000 residents $\theta = 0.001$, $m = 1/\theta = 1000$.

Then $k = \sqrt[3]{m} = 10$, and $n = m^{\frac{2}{3}} = 100$. For a population $P$ of 10 million or $10 \times 10^6$, the total number of tests using Method 2 is given by Eq. (11):

$$L = \dfrac{3P}{m^{2/3}} = \dfrac{3 \times 10 \times 10^6}{100} = 0.3 \times 10^6.$$

Thus the minimum number of tests required using Method 2 is only 0.03 of the total number of tests without grouping test. For comparison, when using Method 1, the minimum number of tests required is equal to $0.63 \times 10^6$, which is 0.063 that without group testing. Method 2 requires less than one-half of the number of tests in Method 1.

In general, $\dfrac{L_{2min}}{L_{1min}} = \dfrac{3}{2} m^{-\frac{1}{6}} = \dfrac{3}{2}\theta^{\frac{1}{6}}. \qquad (12)$
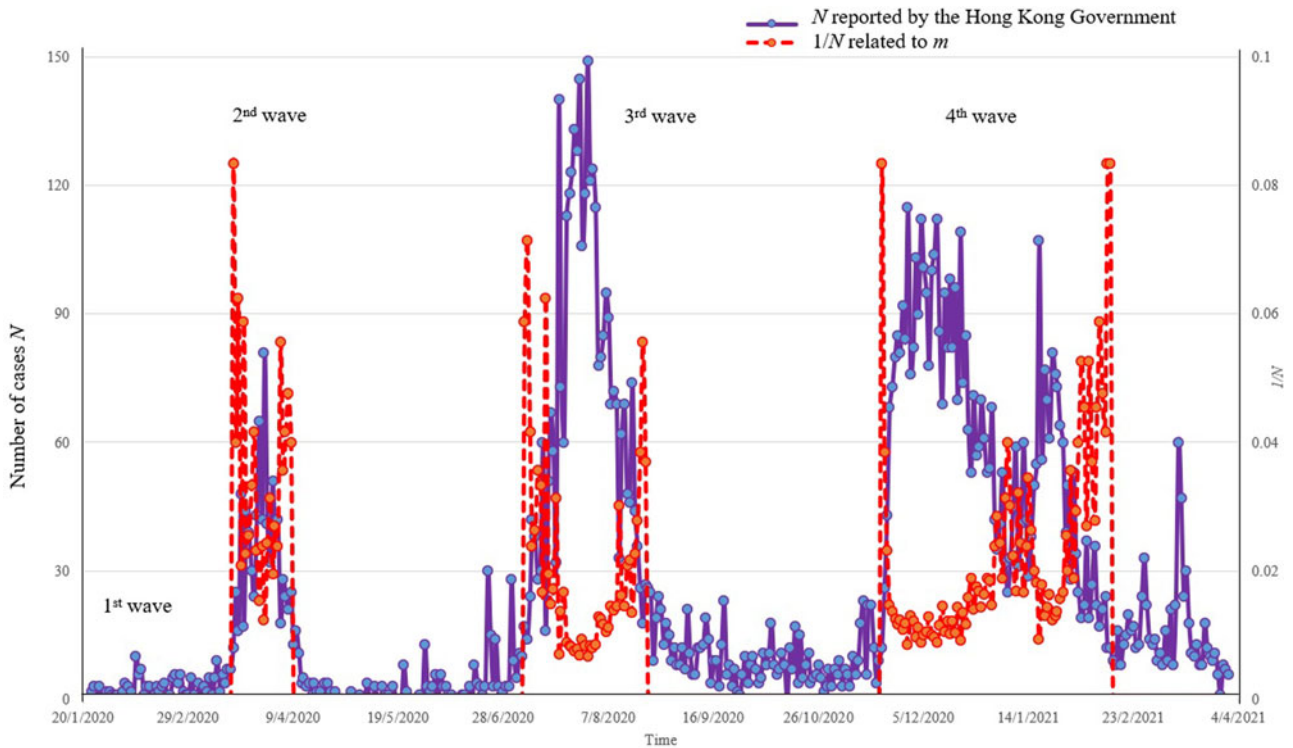
**Fig. 2.** Number of confirmed cases reported by the Hong Kong Government.

As reported by the Hong Kong Special Administrative Region Government [16, 24], there are four waves of infection up to early April 2021 since early 2020. As shown in Figure 2 on the number of daily infection cases $N$ in Hong Kong [24], the infection rate $\theta$ (related to $N$) could be regarded as 150 in 7 million people (or 0.0002) in the second wave. As $m$ is related to $1/\theta$ and $\theta$ is $N/7 \times 10^6$, $1/N$ is also plotted in each wave of infection in Figure 2.

Equation (12) clearly shows that the advantage of Method 2 over Method 1 in reducing the number of tests becomes more prominent as the population infection rate decreases. That is to say, the saving in resources is larger when the infection rate is low. The dependence of ratio $L_{min}/L_o$, where $L_o$ is the number of tests without pooling, on the infection rate $\theta$ is shown in Figure 3 for Method 1 and 2. The comparison of $L_{min}$ for Methods 1 and 2 is shown in Figure 4. The curves in these figures clearly show the advantage of using Method 2 over Method 1, especially for small population infection rate $\theta$.

Note that in practical cases, $\sqrt[3]{m}$ and $m^{\frac{2}{3}}$ may not be integers, and the integers nearest to $\sqrt[3]{m}$ and $m^{\frac{2}{3}}$ are to be chosen as $k$ and $n$.

It should also be pointed out that in the optimisation above, only the mathematical aspect is considered. In practice, other aspects have to be considered. For example, as the group size increases, the reliability of the test results might be compromised due to dilution, and this will increase the likelihood of false-negative test results [8], especially for carriers with low virus load. A balance must be made between increasing the group size and retaining test reliability. In this respect, FDA has provided detailed guidelines for employing pooled sample testing for COVID-19 [25]. The pooling test method is effective and efficient when the positivity rate is low [25].

The robustness of a pooling test method is closely related to the sensitivity and specificity of the test used. In this respect, there is a maximum allowable pool size $n_{max}$ below which the
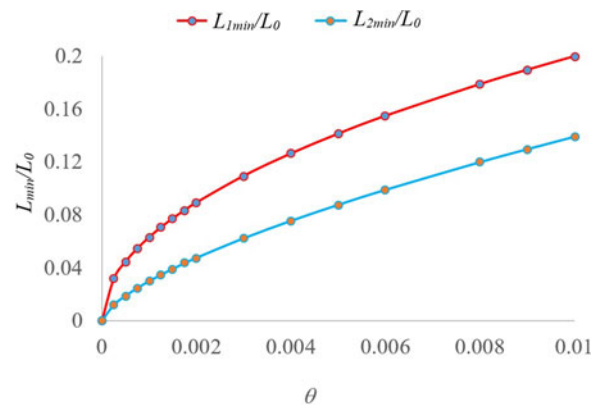


**Fig. 3.** Minimum number of tests required in the two pooling tests relative to that without pooling.
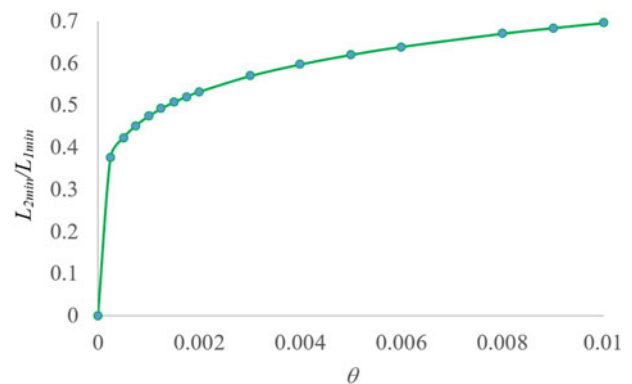


**Fig. 4.** Comparison of the two pooling tests for different infection rates $\theta$.

pooling test may be considered as robust [26]. Thus to achieve robustness, the optimal pool size to be used should be determined, not only from mathematical optimisation, but in conjunction with the limiting value imposed by $n_{max}$. To minimise the influence arising from the uncertainty in prevalence or infection rate, Bish *et al.* [26] also developed a robust pooling strategy via updating pool sizes each week, for each risk group, based on prior week's testing data. This is a more sophisticated approach, though more complex in implementation.

While pooling test is an effective strategy for identifying asymptomatic carriers in mass screening test, symptomatic patients should be dealt with in a different manner in terms of testing for confirmation. In this respect, the point-of-care-testing strategy has been proposed for saving time relative to mass pooling test [14, 27]. Point-of-care (POC) tests are diagnostic tests performed at or near the place where a specimen is collected, such as in hospitals or clinics. Thus transport time of samples to laboratory is minimal, the test is individual, and the symptomatic patient is treated on the spot (POC). Of course, as *a priori* of this strategy, there should be qualified personnel to initially identify the suspected patients before POC testing.

## Conclusions

Pooling tests are commonly used in screening virus carriers. The size of pool is an important parameter as it determines the total number of tests required. A proper choice of group size will yield the minimum number of tests required. Similar to other approaches and methods, there are limitations of the present approach of optimisation and of pooling test methods in general.

### Limitation 1

The optimal group size in pooling test depends on a number of factors, and the present study only considers optimisation with respect to test numbers. Other factors that could affect the optimal group size include the kind of test used and the sensitivity of the test. Thus when considering the optimal group size in pooling test, the results of the present study have to be considered in conjunction with factors such as test sensitivity.

### Limitation 2

As test sensitivity increases, the maximum allowable group size increases. In this paper, we do not consider the upper limit of group size (or equivalently virus concentration in pool) imposed by test sensitivity.

### Limitation 3

Pooling test, be it Method 1 or Method 2, is efficient only when the positivity rate is low because in that case most of the tested cases would be negative and individual testing is not worthwhile and when a large number of people have to be tested in a short time.

### Limitation 4

Pooling test requires two or three rounds of tests, thus requiring a tighter control of logistics to minimise sample handling errors.

### Limitation 5

The results in the present analysis depend on the knowledge of the prevailing infection rate, which is usually changing with time and also not exactly known at the time of test design. An error in *m* will lead to inaccuracy in estimating the minimum number of tests required and hence in testing implementation.

For the two common pooling methods (Method 1, a two-round method, and Method 2, a three-round method), the minimum numbers of tests required are compared in this paper. It is shown that Method 2 is more efficient than Method 1 in reducing the total number of tests, and the reduction increases as the population infection rate decreases. This means that Method 2 requires less cost and resources in terms of test kits and manpower. However, as Method 2 requires three rounds of test, it is more demanding on well-trained personnel in controlling the logistics of the samples. Thus Method 2 should be considered only when a strong team of trained staff members is available.

**Conflict of interest.** None.

**Data availability statement.** Data are available from the authors upon request.

## References

1. **Stephenson J** (2021) Unequal access to COVID-19 vaccines leaves less-wealthy countries more vulnerable, poses threat to global immunity. *JAMA Health Forum* **2**, e210505.
2. **Kayo T** (2020) Global solidarity is necessary to end the COVID-19 pandemic. *Asia-Pacific Review* **27**, 46–56.
3. **Shams SA, Haleem A and Javaid M** (2020) Analyzing COVID-19 pandemic for unequal distribution of tests, identified cases, deaths, and fatality rates in the top 18 countries. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **14**, 953–961.
4. **Haleem A, Javaid ME and Vaishya R** (2020) Effects of COVID-19 pandemic in daily life. *Current Medicine Research and Practice* **10**, 78–79.
5. **Our World in Data**. Statistics and research, coronavirus pandemic (COVID-19) vaccinations, 12 June 2021. Available at https://ourworldindata.org/coronavirus.
6. **Dorfman R** (1943) The detection of defective members of large populations. *The Annals of Mathematical Statistics* **14**, 436–440.
7. **Eberhardt JN, Breuckmann NP and Eberhardt CS** (2020) Multi-stage group testing improves efficiency of large-scale COVID-19 screening. *Journal of Clinical Virology* **128**, 104382.
8. **Mutesa L** *et al.* (2021) A pooled testing strategy for identifying SARS-CoV-2 at low prevalence. *Nature* **589**, 276–280.
9. **Mallapaty S** (2020) The mathematical strategy that could transform coronavirus testing. *Nature* **583**, 504–505.
10. **Hirotsu Y** *et al.* (2020) Pooling RT-qPCR testing for SARS-CoV-2 in 1000 individuals of healthy and infection-suspected patients. *Scientific Reports* **10**, 18899.
11. **Sunjaya AF and Sunjaya AP** (2020) Pooled testing for expanding COVID-19 mass surveillance. *Disaster Medicine and Public Health Preparedness* **14**, E42–E43.
12. **Chow WK and Chow CL** (2020) Response to Sunjaya AF, Sunjaya AP 'pooled testing for expanding COVID-19 mass surveillance'. *Disaster Medicine and Public Health Preparedness* **19**, 1. doi: 10.1017/dmp.2020.454.
13. **Chow WK and Chow CL** (2021) A discussion on implementing pooling detection tests of novel coronavirus (SARS-CoV-2) for large population. *Epidemiology and Infection* **149**, E17.

14. **Chow WK and Chow CL** (2021) Response to Letter to Editor by Zavascki A.P. Urgent need for evaluating point-of-care tests as a RT-PCR-sparing strategy for the diagnosis of Covid-19 in symptomatic patients. *Epidemiology and Infection* **149**, e33.

15. **Wei C** *et al.* (2020) Correlation of population mortality of COVID-19 and testing coverage: a comparison among 36 OECD countries. *Epidemiology and Infection* **28**, 1–3.

16. **Cheung E, Cheung T and Lam J** ( 2021) *Hong Kong fourth wave: mandatory testing of residents in about 70 more buildings, as 56 new coronavirus cases logged*Health & Environment Hong Kong fourth wave: mandatory testing of residents in about 70 more buildings, as 56 new coronavirus cases logged; 19 January.

17. **Abdalhamid B** *et al.* (2020) Assessment of specimen pooling to conserve SARS CoV-2 testing resources. *American Journal of Clinical Pathology* **153**, 715–718.

18. **Eis-Huebinger AM** *et al.* (2020) Ad hoc laboratory-based surveillance of SARS-CoV-2 by real-time RT-PCR using minipools of RNA prepared from routine respiratory samples. *MedRxiv* **127**, 1. doi: 10.1016/j.jcv.2020.104381.

19. **Torres I, Albert E and Navarro D** (2020) Pooling of nasopharyngeal swab specimens for SARS-CoV-2 detection by RT-PCR. *Journal of Medical Virology* **92**, 2306–2307.

20. **Yelin I** *et al.* (2020) Evaluation of COVID-19 RT-qPCR test in multisample pools. *Clinical Infectious Diseases* **71**, 2073–2078.

21. **Pikovski A and Bentele K** (2020) Pooling of coronavirus tests under unknown prevalence. *Epidemiology and Infection* **148**, e183, 1–4.

22. **Grobe N** *et al.* (2021) Sample pooling: burden or solution? *Clinical Microbiology and Infection*. Article in press. doi: 10.1016/j.cmi.2021.04.007.

23. **de Wolff T** *et al.* (2020) Evaluation of pool-based testing approaches to enable population-wide screening for COVID-19. *PLoS ONE* **15**, e0243692.

24. **Data.Gov.HK**. Available at https://data.gov.hk/en-data/dataset/hk-dh-chpsebcddr-novel-infectious-agent/resource/837ebfa2-4e6a-4a28-b442-37c74a7b525e (Accessed 7 April 2021).

25. **Pooled Sample Testing and Screening Testing for COVID-19**. Available at https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/pooled-sample-testing-and-screening-testing-covid-19#pooled.

26. **Bish DR** *et al.* (2021) A robust pooled testing approach to expand COVID-19 screening capacity. *PLoS ONE* **16**, e0246285.

27. **Zavascki AP** (2021) Urgent need for evaluating point-of-care tests as a RT-PCR-sparing strategy for the diagnosis of Covid-19 in symptomatic patients. *Epidemiology and Infection* **149**, e35.