Optimal test pooling for efficient PCR testing of SARS-CoV2

Edgar D. Klenske

Received: 31 May 2020 / Accepted: 3 June 2020 / Published online: 18 September 2020
© Royal Academy of Medicine in Ireland 2020

Letter to the Editor

The new coronavirus SARS-CoV2 is spreading rapidly. Due to its high infectiousness for both pre-symptomatic and asymptomatic cases, it is crucial to test many people quickly. Once infected individuals are identified, they can be quarantined to slow down the spread of the virus. However, the testing capacities are limited and in many regions of the world the demand is already exceeding the existing capacity. Increasing this capacity in normal ways can be difficult due to the availability of test kits and trained laboratory staff.

Recently, several research groups have proposed and developed testing protocols based on test pooling to make better use of the existing testing capacity (see, e.g., [1], [2]). With test pooling, the sample swabs of test subjects are used twice: once in an archive tube and once in the pool tube. Several samples are combined in the pool tube, so that they can be tested simultaneously. The content of the pool tube is analyzed with a test based on the polymerase chain reaction (PCR). Due to the high sensitivity of the PCR test, the result will be positive if at least one of the samples is positive for SARS-CoV2. If the sample pool tests positive, each sample has to be re-tested individually.

For small prior infection probabilities, such a testing procedure can be highly efficient, because many samples can be tested simultaneously with a small number of test kits. Similar testing strategies have formerly successfully been used in sample screening for human immunodeficiency viruses (HIV) [3], influenza [4], and other diseases.

While media outlets report optimistically on the tremendous increase of testing capacity due to pooling methods, it is important to take the prior infection probability into account. If the chosen pool size is too small, more samples could be tested with the same number of test kits. If the chosen pool size is too large, testing capacity will be wasted due to the necessary re-testing, because too many pools are tested positive. In order to make the best use of the existing testing capacity, it is important to select a pool size that is optimal, or at least near-optimal, given the circumstances.

We focus on testing protocols that do not change the sensitivity of the original test [2], because they are especially easy to carry out in the lab and thus less error-prone than tests where varying sensitivity has to be taken into account.

Let $p(C)$ be the prior probability for the samples being positive for SARS-CoV2. This is essentially the prevalence of SARS-CoV2 among the subjects being tested. This probability can be estimated by expert knowledge or field experience, e.g., by analyzing infection numbers of similar cohorts. Since we have to re-test all pooled samples if at least one of them is positive, we are interested in the probability of all the samples being negative $\left(1 - p(C)\right)^{N_P}$, where $N_P$ is the pool size. Under the assumption that the test pooling does not affect the sensitivity of the PCR test, we can now calculate the average number of tests ($N_T$) we have to carry out, based on the prior infection probability and the pool size:

$$N_T(p(C), N_P) = \frac{1}{N_P} \left( \frac{1}{1 + N_P} - \left(1 - p(C)\right)^{N_P} \right)$$

Here we take the one test we always have to carry out and add the number of additional tests for re-testing, which is equivalent to the pool size $N_P$. This number is multiplied with the re-testing probability, the probability of at least one of the samples being positive. Division by the pool size returns the desired result: the expected number of tests per subject.

After all, the average number of tests to determine the infection status of one subject is the number we want to keep low.
The average number of tests over the prevalence, for different pool sizes. Note that for a pool size of \( N_P = 1 \), no re-testing is necessary and \( N_T \) is set to 1 instead of calculating it via Eq. 1.

Low in order to be able to test as many subjects as possible with a given number of test kits. Therefore, we seek to minimize \( N_T \) in Eq. 1. While the prevalence \( p(C) \) is fixed, we can vary the pool size \( N_P \) to achieve this. The average test numbers for different pool sizes over the prior infection probability are shown in Fig. 1. This shows two important aspects: (1) Above a prior infection level of around 30%, no pooling should be done. (2) A pool size of 2 is never optimal and thus should be avoided.

Table 1 lists test pool sizes up to 10 and the prevalence range for which they are optimal. Interestingly, already the prevalence of 3.59% that was measured in the German town Gangelt [5] is high enough to keep the optimal pool size below 10. Since tests are often only carried out only for people with severe symptoms or high-risk contacts of known infected people, a relatively high prior infection probability has to be expected. Overall, the prior infection probability strongly depends on the circumstances. The prior infection probability should therefore be estimated by the medical staff taking the samples, or by the labs themselves, which can track the infection ratio of the samples they have been testing recently.

Even though estimating the prior infection probability precisely can be difficult, using pool testing with a good pool size can be beneficial for expanding the test capacities. Even a simple strategy can already increase the testing capacity a lot, without adding much complexity to the process. One such strategy could be to test with pool size 3 for prevalences up to 30%, and without pooling above 30%. Of course, where ever possible a precise estimate of the prevalence should be used in combination with the optimal pool size.

**Acknowledgments**
The author wants to thank Verena Neumann at Bosch Healthcare Solutions for the valuable discussions, and for proof-reading this letter.

**Data availability**
Not applicable.

**Code availability**
The code can be made available to the reviewers upon request.

**Compliance with ethical standards**

**Conflict of interest**
The author declares no conflict of interest.

**Ethics approval**
Not applicable.

**Consent to participate**
Not applicable.

**References**

1. Pooling method for accelerated testing of COVID-19 (2020). [https://www.technion.ac.il/en/2020/03/pooling-method-for-accelerated-testing-of-covid-19](https://www.technion.ac.il/en/2020/03/pooling-method-for-accelerated-testing-of-covid-19)
2. Corona (2020) Pool testing increases worldwide capacities many times over. [https://healthcare-in-europe.com/en/news/corona-pool-testing-increases-worldwide-capacities-many-times-over.html](https://healthcare-in-europe.com/en/news/corona-pool-testing-increases-worldwide-capacities-many-times-over.html)
3. Soroka S. D. et al (2003) The use of simple, rapid tests to detect antibodies to human immunodeficiency virus types 1 and 2 in pooled serum specimens. J. Clin. Virol. 27:90–96
4. Gaidet N. et al (2012) Saving resources: avian influenza surveillance using pooled swab samples and reduced reaction volumes in real-time RT-PCR. J. Virol. Methods 186:119–125
5. Streeck H. et al (2020) Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event preprint

**Publisher’s note**
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.