An Accurate Model for SARS-CoV-2 Pooled RT-PCR Test Errors

Yair Daon[1,2], Amit Huppert[1,3,*], and Uri Obolski[1,2,*]

[1] *School of Public Health, The Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel*
[2] *Porter School of the Environment and Earth Sciences, The Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel*
[3] *The Biostatistics and Biomathematics Unit, The Gertner Institute for Epidemiology and Health Policy Research, Sheba Medical Center, Tel Hashomer, 52621 Ramat Gan, Israel*
[*] *Equal contribution*

Abstract

PCR testing is an important tool to mitigate outbreaks of infectious diseases. One way of increasing testing throughput is by simultaneously testing multiple samples for the presence of a pathogen, a technique known as *pooling*. During the current COVID-19 pandemic, rapidly testing individuals for the presence of SARS-CoV-2 is conducted in large amounts. Since testing is often a bottleneck in mitigating the spread of SARS-CoV-2, pooling is increasing in popularity. Most analyses of the error rates of pooling schemes assume that including more than a single infected sample in a pooled test does not increase the probability of a positive outcome. We challenge this assumption with experimental data and suggest a novel probabilistic model for the outcomes of pooled tests. As an application, we analyze the false-negative rates of one common pooling scheme known as Dorfman pooling. We show that the false-negative rates of Dorfman pooling increase when the prevalence of infection decreases. However, low infection prevalence is exactly the condition under which Dorfman pooling achieves highest throughput. We therefore implore the cautious use of pooling and development of pooling schemes that consider correctly accounting for tests' error rates.

Introduction

Reverse transcription polymerase chain reaction (RT-PCR) testing is a key component in breaking transmission chains and mitigating the COVID-19 pandemic. As such, the need for large-scale testing has resulted in the development of techniques to increase the the throughput of RT-PCR tests [2, 8, 13, 14, 16]. These techniques, often termed *pooling*, or group testing, have originated in the seminal work of Dorfman and his pooling technique [2, 7]. Under Dorfman pooling, one selects N individuals and performs a single RT-PCR test on their combined (*pooled*) samples. If the pooled test yields a positive result, then each individual is retested separately; otherwise, everyone is declared negative. The throughput efficiency of Dorfman pooling has been demonstrated empirically [2] and its error rates thoroguhly investigated [1, 11, 17].

Studies focused on pooling for SARS-CoV-2 are in a consensus that sample dilution effects [10,21] are not a concern, even for pools as large as 64 individuals [2, 9, 14, 24]. Consequently, studies assumed that the probability of a true-positive (the test's *sensitivity*) does not depend on the number of infected samples in the pool, but rather on the existence of at least one such sample. Thus, the probability of a positive result in a pooled test has been (assumed) identical for a pool with one sample from an infected individual and, e.g., five such samples. This assumption is common in the group testing literature [1, 11], as well as in more specific, COVID-19 focused studies [3, 17]. In this study we challenge this commonly made assumption and show how using a more accurate probabilistic model affects estimation of false-negative rates for Dorfman pooling.

Methods

Formally, we consider a pool containing N individuals $\{1, \ldots, N\}$. We denote the true infection state $\theta \in \{0, 1\}^N$, so individual i is infected iff $\theta_i = 1$. The RT-PCR test's sensitivity (true-positive rate) is denoted S_{e}, and the test's specificity (true-negative rate) is denoted S_{p}. Pooled test result (data) is denoted $\mathbf{d} \in \{0, 1\}$, where $\mathbf{d} = 0$ iff the test returned a negative result.

The common assumption

Previous studies of pooling schemes assumed that the false-negative probability does not depend on the number of infected samples, but merely on the existence of at least one such sample in a pool [1, 11]. Current studies of pooling in the context of SARS-CoV-2 also employ a similar assumption [3, 17]. Explicitly, these studies assume:

$$\mathbb{P}(\mathbf{d} = 0|\theta) = \begin{cases} 1 - S_{\mathrm{e}} & \exists i \text{ such that } \theta_i = 1 \\ S_{\mathrm{p}} & \text{otherwise} \end{cases} \tag{1}$$

Below, we refer to (1) as the *common assumption*.

Refuting the common assumption

We refute the common assumption with experimental data collected from [6], and summarized in Table 1. There, the authors investigate Dorfman pooling and, regardless of the pooled test result, follow up and test each pool member separately. We focus on 128 pools for which at least one

subsequent separate test was positive — of which 29 pooled tests were negative and 99 positive. In the data cited in [6], of the 29 negative pools, subsequent separate testing yielded a single positive result in 24. In contrast, of the 99 positive pools, 42 yielded a single positive test upon subsequent separate testing.

The data in Table 1 allows us to test the following null hypothesis H_0: The probability of a pooled false-negative is equal for pools with one subsequent positively tested member and pools with two or more such members. H_0 is a direct consequence of the common assumption, and rejecting H_0 implies the common assumption is not realistic, at least for SARS-CoV-2.

We apply Fisher's exact test for the presence of more than one positive individual in correctly identified pools. Fisher's test yields an increased odds ratio of 6.4, 95% CI (2.2,23.4), with a p-value $\approx 10^{-4}$. Thus, we reject H_0, refuting the common assumption.

	Negative pool	Positive pool
# subsequent positives = 1	24	42
# subsequent positives ≥ 1	5	57
Total	29	99

Table 1: Contingency table of data from [6]

Our model

Since the essence of the refuted common assumption is that amplification of all samples occurs only once, we assume a more realistic model: amplification of viral RNA succeeds or fails for each sample independently. Furthermore, according to [1,3,11,17], a false-positive does not depend on the number of negative samples in a pool. For lack of data pointing otherwise, we incorporate this assumption into our model with a small modification. We do assume that a false amplification can occur only once per pool. However, we also assume false amplification is independent of any other correct amplification. Specifically, it is possible that every correct amplification fails *and* an erroneous one occurs simultaneously. This assumption is somewhat specific for the current application of screening for SARS-CoV-2 via RT-PCR. For example, cross-reactivity with other coronaviruses would have violated this assumption. However, cross-reactivity was ruled out in [19]. These assumptions lead to the following model, which is illustrated in Figure 1 and summarized in (2).

$$\mathbb{P}(\mathbf{d} = 0|\theta) = S_{\mathrm{p}} \prod_{i=1}^{N} (1 - S_{\mathrm{e}})^{\theta_i} = S_{\mathrm{p}} (1 - S_{\mathrm{e}})^{\sum_i \theta_i}. \tag{2}$$

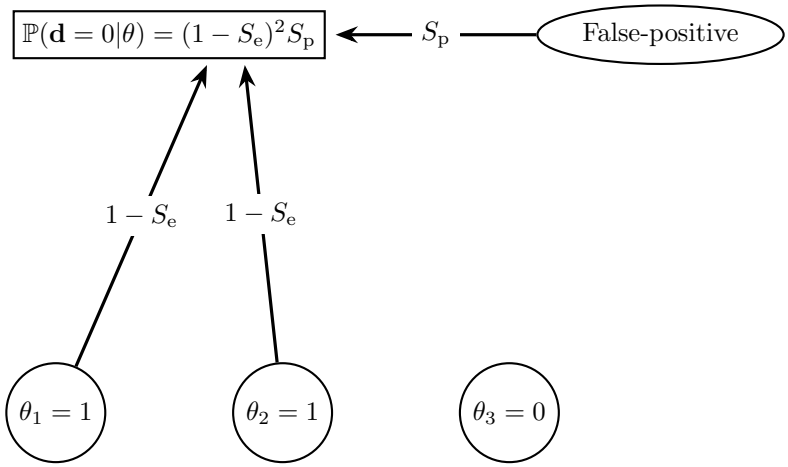


Figure 1: Illustration of the revised probabilistic model. A pool contains individuals $\{1, 2, 3\}$ with state $\theta = (1, 1, 0)$ (i.e., individuals 1 and 2 are infected). A negative pooled test ($\mathbf{d} = 0$) occurs when three detection paths fail: A false-negative occurs for individuals 1 and 2, each with probability $(1 - S_e)$. Additionally, no false-positive detection occurs, with probability S_p. Individual 3 is not infected and does not contribute to the probability of the pooled test result.

Application: scheme false-negative rate

We calculate the false-negative rate for a single *infected* individual, henceforth referred to as "Donald", under our model and under the common assumption. We distinguish three types of false-negative events when performing pooling. A *single test*'s false-negative is the event of a negative result upon testing Donald separately, i.e., in an RT-PCR test without pooling. A *pooled* false-negative occurs when a pooled test containing Donald's sample (and other samples) yields a negative result, i.e., the pooling fails to detect at least one positive result. Lastly, a *scheme* false-negative occurs when an entire pooling scheme fails to identify Donald as infected. Our first task is to calculate Dorfman's scheme false-negative rate. Equivalently, we ask: what is the probability of not identifying Donald as infected under Dorfman pooling?

Denote the prevalence of infection in the (tested) population q. We denote Donald as individual 1, so that $\theta_1 = 1$. Then:

$$
\begin{aligned}
\mathbb{P}(\mathbf{d}=0|\theta_1=1) &= \sum_{\theta_2,\ldots,\theta_N} \mathbb{P}(\theta_2,\ldots,\theta_N)\mathbb{P}(\mathbf{d}=0|\theta_1=1,\theta_2,\ldots,\theta_N) \\
&= \sum_{k=0}^{N-1} \mathbb{P}(\sum_{i=2}^{N}\theta_i=k)\mathbb{P}(\mathbf{d}=0|\theta_1=1,\sum_{i=2}^{N}\theta_i=k) \\
&= \sum_{k=0}^{N-1} \binom{N-1}{k} q^k(1-q)^{N-1-k} S_{\mathrm{p}}(1-S_{\mathrm{e}})^{1+k} \\
&= S_{\mathrm{p}}(1-S_{\mathrm{e}}) \sum_{k=0}^{N-1} \binom{N-1}{k} \left(q(1-S_{\mathrm{e}})\right)^k (1-q)^{N-1-k} \\
&= S_{\mathrm{p}}(1-S_{\mathrm{e}})(1-qS_{\mathrm{e}})^{N-1}.
\end{aligned} \tag{3}
$$

If the pooled test yields a positive result, Donald is tested separately. Taking a conservative stand, it is assumed that such a simple procedure poses no risk of introducing contaminant RNA. Therefore, the separate test yields a positive result with probability S_{e}.

We calculate the probability that Donald is mistakenly identified as not infected, henceforth referred to as the scheme's false-negative rate and denoted S_{fn}. In order to correctly identify an infected individual as infected, both pooled and separate tests have to yield a positive result. Thus, the scheme's false-negative rate is:

$$
\begin{aligned}
S_{\mathrm{fn}} :&= 1 - S_{\mathrm{e}}\mathbb{P}(\mathbf{d}=1|\theta_1=1) \\
&= 1 - S_{\mathrm{e}}\left[1 - S_{\mathrm{p}}(1-S_{\mathrm{e}})(1-qS_{\mathrm{e}}))^{N-1}\right].
\end{aligned} \tag{4}
$$

Comparison metric

The single test false-negative rate $1 - S_{\mathrm{e}}$ and scheme false-negative rate S_{fn} are compared via:

$$
E_{\mathrm{exact}} := \frac{S_{\mathrm{fn}} - (1-S_{\mathrm{e}})}{1-S_{\mathrm{e}}} \cdot 100. \tag{5}
$$

E_{exact} is the percentage increase in the pooling scheme false-negative rate, relative to the single test false-negative rate.

According to the common assumption, the scheme false-negative rate is $1 - S_{\mathrm{e}}^2$. A straight forward calculation shows that this implies the percentage increase in scheme false-negative rate is $E_{\mathrm{common}} := 100 \cdot S_{\mathrm{e}}$.

Results

Scheme false-negative

We plot E_{exact} for varying prevalence q and sensitivity S_{e} values, and make the comparison with E_{common}. As recommended by [2], we apply different pool sizes N, for different prevalence values. We observe that for a false-positive rate $S_{\mathrm{p}} = 0.95$ [2] and a range of reasonable sensitivity and prevalence values [19, 20, 22, 23], an increase of at least 60% in E_{exact} can be expected (Figure 2).

Interestingly, an increase in infection prevalence monotonically decreases the scheme false-negative rate, as can also be easily seen from (4). For the chosen parameter ranges, the increase in the single test false-negative rates increases the relative error E_{exact}. These effects can be seen in Figure 2 (left panel), upon conditioning on pool size. Extending the range for S_{p} yields no qualitative differences. We further compare E_{exact} to E_{common}, showing the discrepancy changes as a function of both prevalence and the single test sensitivity (Figure 2, right panel).

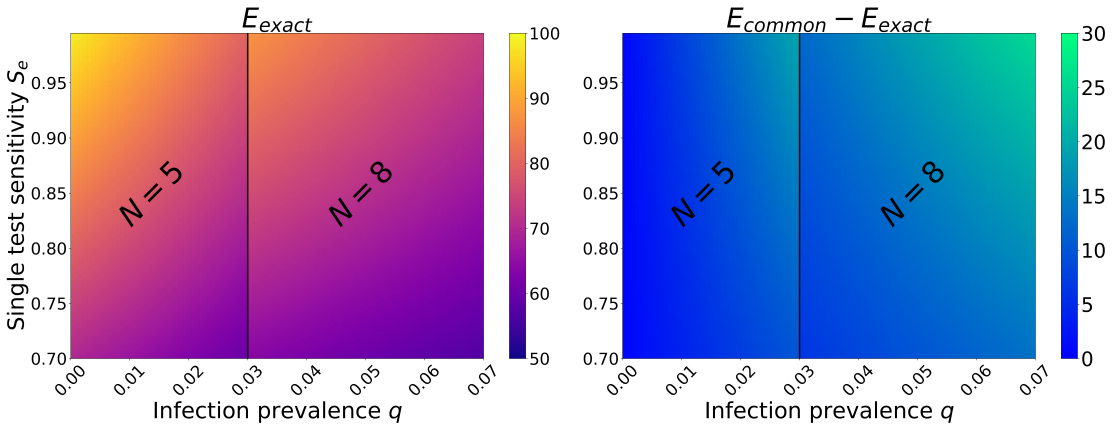


Figure 2: Relative increase in Dorfman pooling false-negative rates E_{exact}. Left: Colors represent E_{exact}, the relative percentage increase in the scheme false-negative rates relative to the single test false-negative rates (5). Right: colors represent the difference between E_{common} and E_{exact}. The disease prevalence q, is varied on the x-axis, while the test sensitivity is varied on the y-axis. Pool size N, was chosen according to q as in [2]. The left panel shows that E_{exact} is largest for low prevalence values q. The difference between E_{common} and E_{exact} can be as large as 30%, as seen in the right panel.

Discussion

Here, we developed a more realistic and novel probabilistic model for the outcomes of pooled PCR tests, parametrized for SARS-CoV-2. Contrary to the common assumption, we assume, based on data (Table 1), that multiple infected individuals increase the likelihood of a positive pooled test result. A direct consequence of our model is that false-negative rates depend on infection prevalence. Specifically, low values of infection prevalence increase the false-negative rates of Dorfman pooling. These results remain qualitatively similar under varying parameter values, in the observed ranges [12, 19, 20, 22] (Figure 2). Our results give rise to a conflict: low infection prevalence leads to high efficiency of Dorfman pooling [2], while also increasing false-negative rates.

As the COVID-19 pandemic progresses, the infection prevalence in various tested populations undergoes frequent changes. Hence, as our results suggest, pooling schemes employed for mass testing should be used with caution in populations where infection rates are low. Such mass-tested populations often include air travel passengers [5] or presymptomatic and asymptomatic individuals [18], and can be crucial for controlling outbreaks [15].

Thus, improving pooling schemes is imperative. An especially important consideration in de-

signing such schemes is explicitly taking the intrinsic PCR error rates into account. We [4], as well as others [16], have developed such models, although they have not yet been implemented in real-world settings.

To conclude, pooling is an important technique that can increase testing throughput in a cost-effective manner. Nevertheless, care must be given to pooling schemes' false-negative rates, especially under low infection prevalence settings.

Acknowledgements

Yair Daon was supported by a post-doctoral fellowship from the Tel Aviv University Center for Combating Pandemics and the Raymond and Beverly Sackler dean's post-doctoral fellowship.

References

[1] Hrayer Aprahamian, Douglas R Bish, and Ebru K Bish, *Optimal group testing: Structural properties and robust solutions, with application to public health screening*, INFORMS Journal on Computing (2020).

[2] Netta Barak, Roni Ben-Ami, Tal Sido, Amir Perri, Aviad Shtoyer, Mila Rivkin, Tamar Licht, Ayelet Peretz, Judith Magenheim, Irit Fogel, et al., *Lessons from applied large-scale pooling of 133,816 SARS-CoV-2 RT-PCR tests*, Science Translational Medicine (2021).

[3] Alhaji Cherif, Nadja Grobe, Xiaoling Wang, and Peter Kotanko, *Simulation of pool testing to identify patients with coronavirus disease 2019 under conditions of limited test availability*, JAMA network open **3** (2020), no. 6, e2013075–e2013075.

[4] Yair Daon, Amit Huppert, and Uri Obolsik, *DOPE: D-Optimal Pooling Experimental design with application to SARS-CoV-2 screening*, arXiv (2021).

[5] Yair Daon, Robin N Thompson, and Uri Obolski, *Estimating covid-19 outbreak risk through air travel*, Journal of Travel Medicine **27** (2020), no. 5, taaa093.

[6] Adolfo de Salazar, Antonio Aguilera, Rocio Trastoy, Ana Fuentes, Juan Carlos Alados, Manuel Causse, Juan Carlos Galán, Antonio Moreno, Matilde Trigo, Mercedes Pérez-Ruiz, et al., *Sample pooling for SARS-CoV-2 RT-PCR screening*, Clinical Microbiology and Infection **26** (2020), no. 12, 1687.e1–1687.e5.

[7] Robert Dorfman, *The detection of defective members of large populations*, The Annals of Mathematical Statistics **14** (1943), no. 4, 436–440.

[8] Rudolf Hanel and Stefan Thurner, *Boosting test-efficiency by pooled testing for SARS-CoV-2—Formula for optimal pool size*, PLoS One **15** (2020), no. 11, e0240652.

[9] Yosuke Hirotsu, Makoto Maejima, Masahiro Shibusawa, Yuki Nagakubo, Kazuhiro Hosaka, Kenji Amemiya, Hitomi Sueki, Miyoko Hayakawa, Hitoshi Mochizuki, Toshiharu Tsutsui, et al., *Pooling RT-qPCR testing for SARS-CoV-2 in 1000 individuals of healthy and infection-suspected patients*, Scientific reports **10** (2020), no. 1, 1–8.

[10] Frank K Hwang, *Group testing with a dilution effect*, Biometrika **63** (1976), no. 3, 671–680.

[11] Hae-Young Kim, Michael G Hudgens, Jonathan M Dreyfuss, Daniel J Westreich, and Christopher D Pilcher, *Comparison of group testing algorithms for case identification in the presence of test errors*, Biometrics **63** (2007), no. 4, 1152–1163.

[12] Lauren M Kucirka, Stephen A Lauer, Oliver Laeyendecker, Denali Boon, and Justin Lessler, *Variation in false-negative rate of reverse transcriptase polymerase chain reaction—based SARS-CoV-2 tests by time since exposure*, Annals of Internal Medicine (2020).

[13] Thomas J Kwiatkowski Jr, Huda Y Zoghbi, Susan A Ledbetter, Kimberly A Ellison, and A Craig Chinault, *Rapid identification of yeast artificial chromosome clones by matrix pooling and crude lysate PCR.*, Nucleic acids research **18** (1990), no. 23, 7191.

[14] Stefan Lohse, Thorsten Pfuhl, Barbara Berkó-Göttel, Jürgen Rissland, Tobias Geißler, Barbara Gärtner, Sören L Becker, Sophie Schneitler, and Sigrun Smola, *Pooling of samples for testing for SARS-CoV-2 in asymptomatic people*, The Lancet Infectious Diseases (2020).

[15] Michael J Mina and Kristian G Andersen, *COVID-19 testing: One size does not fit all*, Science **371** (2021), no. 6525, 126–127.

[16] Rodrigo Noriega and Matthew Samore, *Increasing testing throughput and case detection with a pooled-sample Bayesian approach in the context of COVID-19*, bioRxiv (2020).

[17] Alexander Pikovski and Kajetan Bentele, *Pooling of coronavirus tests under unknown prevalence*, Epidemiology & Infection **148** (2020).

[18] Robin N Thompson, Francesca A Lovell-Read, and Uri Obolski, *Time from symptom onset to hospitalisation of coronavirus disease 2019 (COVID-19) cases: implications for the proportion of transmissions from infectors with few symptoms*, Journal of clinical medicine **9** (2020), no. 5, 1297.

[19] Puck B van Kasteren, Bas van Der Veer, Sharon van den Brink, Lisa Wijsman, Jørgen de Jonge, Annemarie van den Brandt, Richard Molenkamp, Chantal BEM Reusken, and Adam Meijer, *Comparison of seven commercial RT-PCR diagnostic kits for COVID-19*, Journal of Clinical Virology **128** (2020), 104412.

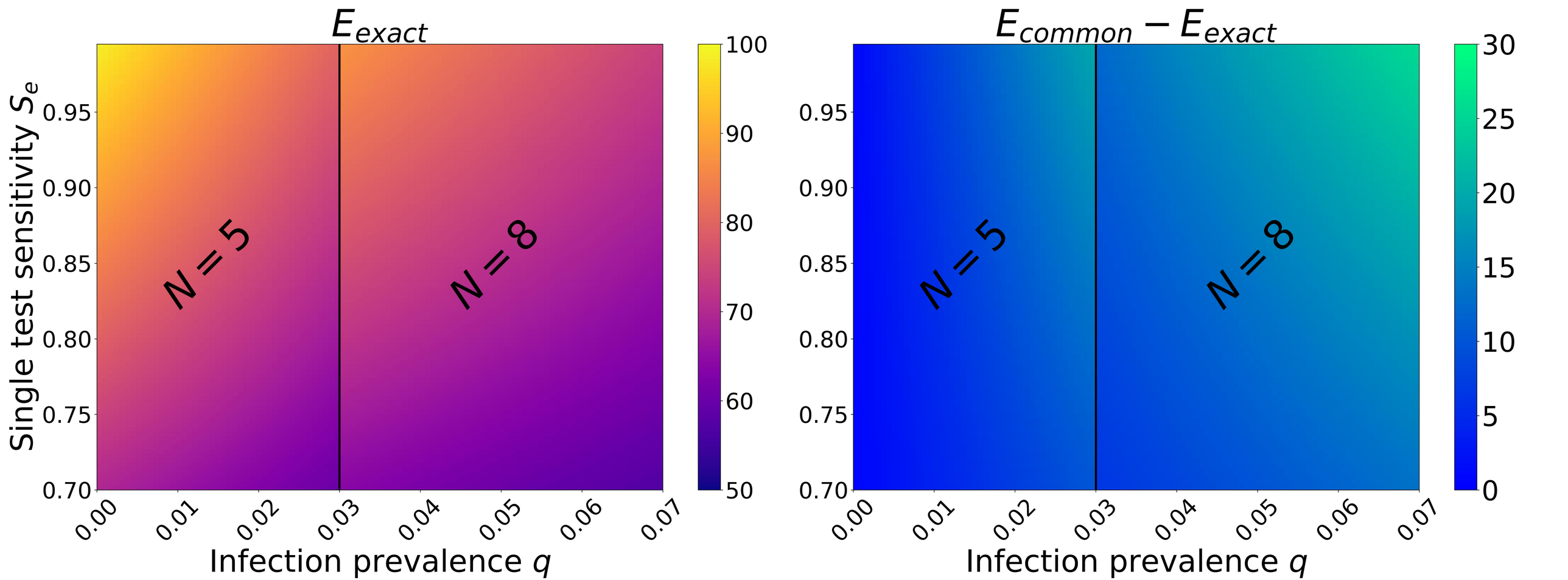
[20] Jessica Watson, Penny F Whiting, and John E Brush, *Interpreting a COVID-19 test result*, BMJ **369** (2020).

[21] Lawrence M Wein and Stefanos A Zenios, *Pooled testing for HIV screening: capturing the dilution effect*, Operations Research **44** (1996), no. 4, 543–569.

[22] Paul S Wikramaratna, Robert S Paton, Mahan Ghafari, and José Lourenço, *Estimating the false-negative test probability of SARS-CoV-2 by RT-PCR*, Eurosurveillance **25** (2020), no. 50, 2000568.

[23] Steven Woloshin, Neeraj Patel, and Aaron S Kesselheim, *False negative tests for SARS-CoV-2 infection-challenges and implications*, New England Journal of Medicine (2020).

[24] Idan Yelin, Noga Aharony, Einat Shaer Tamar, Amir Argoetti, Esther Messer, Dina Berenbaum, Einat Shafran, Areen Kuzli, Nagham Gandali, Omer Shkedi, et al., *Evaluation of covid-19 rt-qpcr test in multi sample pools*, Clinical Infectious Diseases **71** (2020), no. 16, 2073–2078.


E_{exact}
$E_{common} - E_{exact}$
Single test sensitivity S_e
Infection prevalence q
Infection prevalence q
$N = 5$
$N = 8$
$N = 5$
$N = 8$
0.70
0.75
0.80
0.85
0.90
0.95
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
100
90
80
70
60
50
30
25
20
15
10
5
0