medRxiv preprint doi: https://doi.org/10.1101/2020.12.02.20242651; this version posted January 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Inflated false-negative rates in pooled RT-PCR tests of SARS-CoV-2

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

[1] *Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv, Israel*
[2] *School of Public Health, Tel Aviv University, Tel Aviv, Israel*
[3] *The Gertner Institute for Epidemiology and Health Policy Research, Tel Hashomer, Israel*

January 23, 2021

Abstract

Background: Pooling is a popular strategy for increasing SARS-CoV-2 testing throughput. One popular pooling scheme is Dorfman pooling: test N individuals simultaneously. If the test is positive — retest each individual separately. However, requiring more than one positive test may lead to increased false-negative rates.

Methods: We analyze the false-negative rate (i.e., the probability of a negative result for an infected individual) of Dorfman pooling via a new probabilistic model. We demonstrate that different, previously made probabilistic assumptions regarding pooling are unlikely in light of empiric data. Our model is conservative in that it ignores sample dilution effects, which can only worsen pooling performance.

Results: We show that one can expect a 60-80% increase in false-negative rates under Dorfman pooling, for reasonable parameter values. Moreover, we show that the false-negative rates under Dorfman pooling increase when the prevalence of infection decreases.

Discussion: In most pooling schemes, identifying an infected individual requires positive results in multiple tests and hence substantially increases false-negative rates. Furthermore, this phenomenon is more pronounced when infection prevalence is low — exactly when pooling is most efficient. Thus, pooling presents an inherent trade-off: it is most efficient when it is least accurate. The deterioration of false-negative rates and the aforementioned trade-off are inherent problems of pooling schemes and should be kept in mind by practitioners and policy makers.

1 Introduction

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 pooling schemes of RT-PCR tests [2, 6, 7, 10, 11]. One such popular scheme is Dorfman pooling [2, 5]: Select N individuals and perform a single RT-PCR test on their combined ("pooled") samples. If the pooled test yields a positive result — test each individual separately. The throughput efficiency of Dorfman pooling has been demonstrated empirically [2]. However, when test error rates are taken into consideration, a sharp increase in false-negative rates can be expected.

It is important to distinguish three types of false-negative events when performing pooling. For convenience, we follow a single *infected* individual, henceforth referred to as "Donald". 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. We denote the test sensitivity S_e, so the probability of a single test false-negative is $1 - S_e$. 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 goal is to calculate Dorfman's scheme false-negative rate. Rephrasing, we wish to answer the following question: what is the probability of not identifying Donald as infected under a Dorfman pooling scheme?

2 Methods

2.1 Probabilistic Assumptions

We assume two pathways for a positive pooled test result: Viral RNA from an infected individual is correctly detected; or, some erroneous detection occurs (e.g. contaminant viral RNA is introduced). We ignore cross-reactivity with other Coronaviruses, which is negligible [14]. We assume a homogeneous and disconnected population (each individual is infected independently and with equal probability). For simplicity, we do not take into account sample dilution, since it can only further increase false-negative rates [2].

Our assumptions, although natural, stand in contrast to assumptions commonly made in the literature [1, 3, 8, 12]. It is commonly assumed that the probability of a positive pooled test is not increased by having more than one infected individual

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.02.20242651; this version posted January 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

Table 1: Contingency table of data from [4]

in the pool. We refute the common assumptions with experimental data summarized in table 1, collected from [4]. There, the authors investigate Dorfman pooling and, regardless of the pooled test result, follow up and test each pool member separately. Here, 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 [4], 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. 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 independence of a pooled result from one or more individuals infected within the pool, as assumed in [1, 3, 8, 12]

2.2 calculation of Dorfman's scheme false-negative rate

Denote the prevalence of infection in the (tested) population q. As before, S_e is the test's sensitivity, so $1 - S_e$ is the single RT-PCR test's false-negative rate. We also denote S_p the test specificity — the true-negative probability. This encompasses events of an erroneous RNA detection (via any possible pathway), which may cause a false-positive. By our assumptions, a pool containing Donald's sample and $N - 1$ other samples will yield a *negative* result if all of the following occur:

- No erroneous detection occurred (i.e. no false-positive). This happens with probability S_p.
- The detection process fails for Donald's sample. A false-negative occurs for Donald, with probability $1 - S_e$.
- No detection for any of the other $N-1$ samples. For a single sample, the probability of being detected is the prevalence of SARS-CoV-2 in the tested population q, multiplied by the sensitivity S_e, hence the probability of detection is qS_e. For $N - 1$ such samples, the probability of *not* being identified is $(1 - qS_e)^{N-1}$.

The pooled false-negative probability for Donald is simply the product of the terms above. Hence:

$$\begin{aligned}\mathbb{P}(\text{pool is positive}) &= 1 - \mathbb{P}(\text{pool false-negative}) \\ &= 1 - S_p(1 - S_e)(1 - qS_e)^{N-1}.\end{aligned} \tag{1}$$

If the pooled test yields a positive result, Donald is tested separately. We assume 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 — the scheme's false-negative rate — denoted P_{sfn} below. To *correctly* identify Donald as infected, both pooled and separate tests have to yield a positive result. Thus, the scheme's false-negative rate P_{sfn} is the complement of the product of the two previous terms:

$$\begin{aligned}P_{\text{sfn}} :&= 1 - \mathbb{P}(\text{correctly identify Donald as infected}) \\ &= 1 - S_e\left[1 - S_p(1 - S_e)(1 - qS_e))^{N-1}\right].\end{aligned} \tag{2}$$

2.3 Comparison metric

The single test false-negative rate $1 - S_e$ and scheme false-negative rate P_{sfn} are compared via:

$$E_{\text{rel}} := \frac{P_{\text{sfn}} - (1 - S_e)}{1 - S_e} \cdot 100. \tag{3}$$

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

For the approximation commonly used in the literature estimates the scheme false-negative rate as $1 - S_e^2$ (section **??** and [1, 3, 8, 12]). A short calculation shows that this approximation implies the percentage increase in scheme false-negative rate is just $100 \cdot S_e$.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.02.20242651; this version posted January 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

3 Results

We plot E_{rel} for varying prevalence q and sensitivity S_e values. As recommended by [2], we apply different pool sizes N, for different prevalence values. We observe that for a false-positive rate $S_p = 0.95$ [2] and a range of reasonable sensitivity and prevalence values [13–16], an increase of at least 60% in E_{rel} can be expected (Figure 1). Interestingly, an increase in infection prevalence monotonically decreases the scheme false-negative rate, as can also be easily seen from equation (2). For the chosen parameter ranges, the increase in the single test false-negative rates increases the relative error E_{rel}. These effects can be seen in Figure 1 (left panel), upon conditioning on pool size. Extending the range for S_p yields no qualitative differences. We further compare E_{rel} to the commonly used approximation, showing the discrepancy changes as a function of both prevalence and the single test sensitivity (Figure 1, right panel)

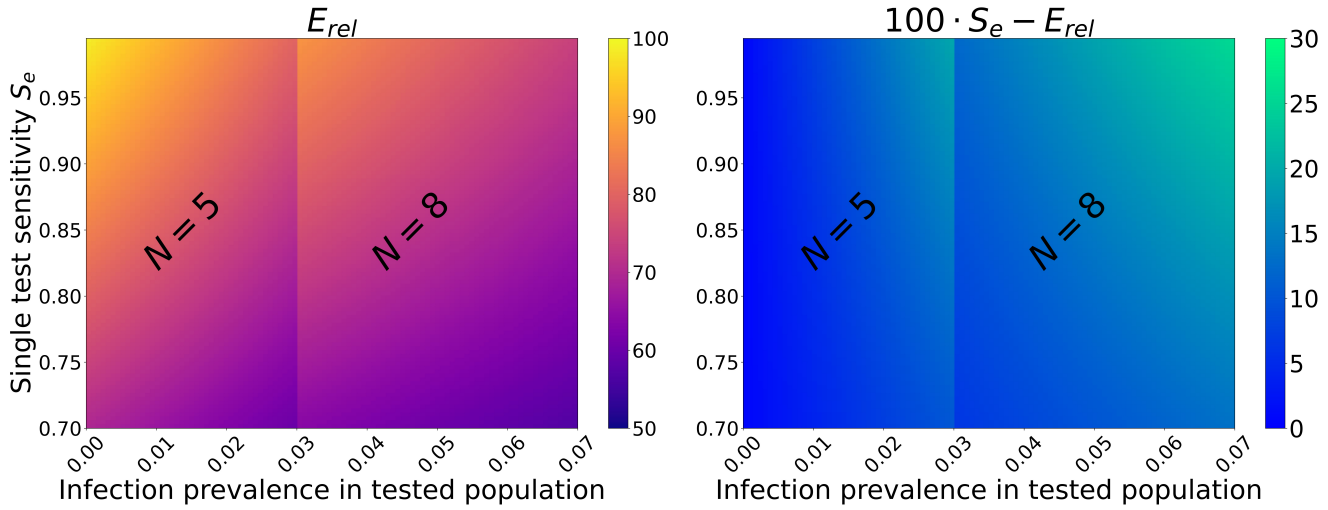


Figure 1: Relative increase in Dorfman pooling false-negative rates E_{rel}. Left: Colors represent E_{rel}, the relative percentage increase in the scheme false-negative rates relative to the single test false-negative rates (eq (3)). Right: colors represent the difference between the approximation assumed in other studies and E_{rel} as calculated by our model. 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].

4 Discussion

Although Dorfman pooling improves testing throughput, we have shown that it can increase E_{rel} — the false-negative rates relative to individual testing. Furthermore, low values of infection prevalence, or low values of single test false-negative rates, increase E_{rel}. These results remain qualitatively similar under varying parameter values, in the observed ranges [9, 13–15] (Figure 1).

Our results lead to a fact almost disregarded in [2]: although (Dorfman) pooling is most efficient when prevalence is low, such circumstances are exactly those leading to a substantial increase in false-negative rates.

An approximation to our results has been previously considered: $P_{\text{sfn}} \approx 1 - S_e^2$ [1, 3, 8, 12]. Such an approximation arises when one assumes that the probability of a false-negative is identical for the pooled and single tests. Although simplistic in nature, this approximation does capture the intuition behind our results: For Donald to be considered negative under Dorfman pooling, he has to test negative twice. In addition to this approximation only being an upper bound [3], it does not account for the effect of infection prevalence on the false-negative rates under Dorfman pooling. Supporting the importance of this effect, such an association between prevalence and the scheme false-negative rate under Dorfman pooling has been empirically noted [2].

Although we have shown the inherent risk of Dorfman pooling, this shortcoming applies to other pooling schemes. Pooling schemes (e.g. [6, 8, 17]), require some sequence of positive pooled results to correctly identify Donald as infected. Consider the hierarchical pooling scheme [8, 17]: If the first pool yields a positive result, it is split in two. Then the splitting is repeated until resulting pools are negative or individuals are tested separately. With an initial pool size of 32, Donald will necessarily have to test positive in pools of size 32, 16, 8, 4 and 2, as well as in a single test, for the scheme to correctly identify him as infected. Compare this to the Dorfman scheme that requires a positive test in a pool of size $N = 8$, and an additional single positive test to identify Donald as infected. The hierarchical pooling scheme of [8, 17] will necessarily yield more false-negatives than Dorfman pooling — there are additional places for it to fail.

As mentioned in [2], introducing a positive dependence within a pool decreases the false-positive rate. In the extreme case, consider a fully connected pool, where one infection implies the entire pool is infected. In this case, a calculation analogous to the one conducted above recovers the initial false-negative rate $1 - S_e$. Interestingly, pooling was also noted to

medRxiv preprint doi: https://doi.org/10.1101/2020.12.02.20242651; this version posted January 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

have increased throughput when infection probabilities are dependent between the pooled individuals [2], providing another advantage to sampling dependent individuals in pooling schemes.

To conclude, pooling is an important technique which can improve testing throughput in a cost-effective manner. Nevertheless, a substantial increase in pooling schemes' false-negative rates can be expected. Such a increase has crucial implications for controlling the spread of COVID-19.

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*, medRxiv (2020).

[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] 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.

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

[6] US Food, Drug Administration, et al., *Accelerated emergency use authorization (eua) summary COVID-19 RT-PCR test (laboratory corporation of america)*.

[7] Rudolf Hanel and Stefan Thurner, *Boosting test-efficiency by pooled testing strategies for SARS-CoV-2*, arXiv preprint arXiv:2003.09944 (2020).

[8] 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 error*, Biometrics **63** (2007), no. 4, 1152–1163.

[9] 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).

[10] 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).

[11] 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).

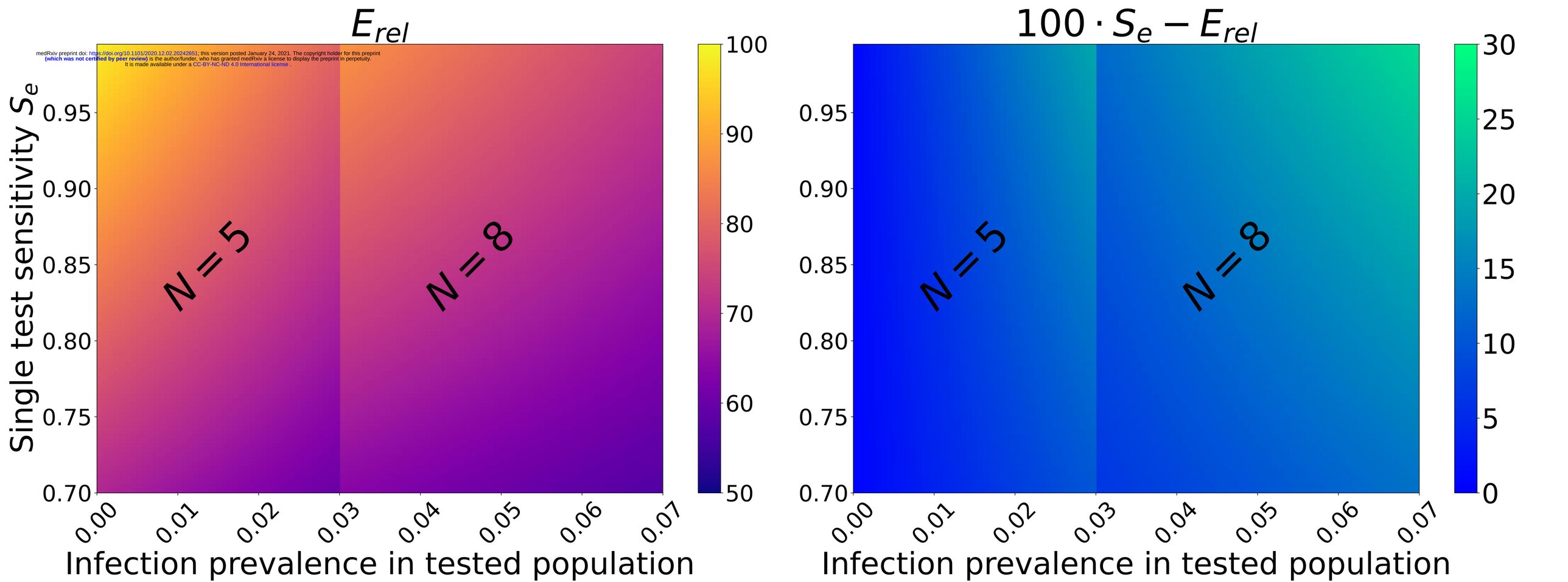
[12] Alexander Pikovski and Kajetan Bentele, *Pooling of coronavirus tests under unknown prevalence*, Epidemiology & Infection **148** (2020).

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

[14] L Wijsman, R Molenkamp, CBEM Reusken, A Meijer, et al., *Comparison of seven commercial RT-PCR diagnostic kits for COVID-19.*, Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology **128** (2020), 104412–104412.

[15] Paul Wikramaratna, Robert S Paton, Mahan Ghafari, and Jose Lourenco, *Estimating false-negative detection rate of SARS-CoV-2 by RT-PCR*, medRxiv (2020).

[16] 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).

[17] Idan Yelin, Noga Aharony, Einat Shaer-Tamar, Amir Argoetti, Esther Messer, a Berenbaum, Einat Shafran, Areen Kuzli, Nagam Gandali, Tamar Hashimshony, et al., *Evaluation of COVID-19 RT-qPCR test in multi-sample pools*, medRxiv (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.02.20242651; this version posted January 24, 2021. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.