medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 .

False positives in reverse transcription PCR testing for SARS-CoV-2

Andrew N. Cohen[1]*, Bruce Kessel[2]

[1] Center for Research on Aquatic Bioinvasions, Richmond CA, USA

[2] Department of Obstetrics, Gynecology and Women's Health, John A. Burns School of Medicine, University of Hawaii, Honolulu HI, USA

* Correspondence: acohen@bioinvasions.com

Abstract

Background

Large-scale testing for SARS-CoV-2 by RT-PCR is a key element of the response to COVID-19, but little attention has been paid to the potential frequency and impacts of false positives.

Methods

From a meta-analysis of external quality assessments of RT-PCR assays of RNA viruses, we derived a conservative estimate of the range of false positive rates that can reasonably be expected in SARS-CoV-2 testing, and analyzed the effect of such rates on analyses of regional test data and estimates of population prevalence and asymptomatic ratio.

Findings

Review of external quality assessments revealed false positive rates of 0-16.7%, with an interquartile range of 0.8-4.0%. Such rates would have large impacts on test data when prevalence is low. Inclusion of such rates significantly alters four published analyses of population prevalence and asymptomatic ratio.

Interpretation

The high false discovery rate that results, when prevalence is low, from false positive rates typical of RT-PCR assays of RNA viruses raises questions about the usefulness of mass testing; and indicates that across a broad range of likely prevalences, positive test results are more likely to be wrong than are negative results, contrary to public health advice about SARS-CoV-2 testing. There are myriad clinical and case management implications. Failure to appreciate the potential frequency of false positives and the consequent unreliability of positive test results across a range of scenarios could unnecessarily remove critical workers from service, expose uninfected individuals to greater risk of infection, delay or impede appropriate medical treatment, lead to inappropriate treatment, degrade patient care, waste personal protective equipment, waste human resources in unnecessary contact tracing, hinder the development of clinical improvements, and weaken clinical trials. Measures to raise awareness of false positives, reduce their frequency, and mitigate their effects should be considered.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

Introduction

Large-scale testing for SARS-CoV-2 is a key element of the response to the COVID-19 pandemic. Test protocols use the reverse transcription polymerase chain reaction (RT-PCR) to detect diagnostic sequences in the RNA genome. The results are used to identify transmission clusters, model rates of spread, calculate death rates, assess the number and significance of asymptomatic carriers, shape public policy on social distancing, and inform decisions about the spatial and temporal allocation of medical resources.

Test accuracy is thus of paramount importance, yet little attention has been paid to the potential frequency and impacts of false positives in SARS-CoV-2 testing. We hypothesize that a significant false positive rate (FPR) is likely and could have substantial implications for clinical care, epidemiologic statistics, and public health policy.

Methods

We searched on-line for (1) external quality assessments (EQAs) of diagnostic laboratories conducting RT-PCR assays for RNA viruses; (2) studies that estimated false negative rates (FNRs) in SARS-CoV-2 testing; and (3) studies that used SARS-CoV-2 RT-PCR test data to estimate population prevalence or the asymptomatic ratio, defined as the ratio of asymptomatic infected individuals—in the sense of permanently asymptomatic, not presymptomatic—to total infected individuals.

We conducted a meta-analysis of EQAs of medical diagnostic laboratories conducting RT-PCR assays for RNA viruses, in order to calculate FPRs for tests similar to the SARS-CoV-2 RT-PCR tests. We excluded EQAs prior to 2004, since many of the assays relied on older RT-PCR methods that may be less accurate. The EQAs provided participating laboratories with blind panels of positive and negative samples to assay, and received and analyzed the results. For each EQA we calculated the FPR as the number of negative samples reported as positive, divided by the total number of negative samples. Where we could only determine a range, we conservatively took the FPR as the lower bound of the range. In EQAs where no negative samples were reported as positive, we report the FPR in Table 1 as below a detection limit equal to the reciprocal of the total number of negative samples, but for analysis treated the FPR as zero. We calculated the median value and interquartile range of these FPR data and also of a subset restricted to EQAs where the total number of negative samples was >100. We used the lower of these medians and interquartile ranges to analyze the effects of false positives. From published studies we determined a range of FNR values to use in this analysis. We obtained test data and calculated test positivity rates for South Korea[1] and New York State,[2] representing regions with low and high positivity rates.

We derived the dependence of the false discovery rate (FDR)—the proportion of positive test results that are false positives—on the FPR, FNR and test positivity rate (Appendix pp.1–2). We modeled the effect of these on FDR across ranges for FPR, FNR and positivity based on, respectively, the EQA analysis, published estimates, and

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

the two sets of regional test data. In studies that we identified as using SARS-CoV-2 RT-PCR test data to estimate population prevalence or the asymptomatic ratio, if FPRs and FNRs were not included in the analysis we re-analyzed the data using FPR and FNR values as described above.

Results

We compiled data and calculated FPRs for 43 EQAs conducted between 2004 and 2019, each of which assessed between three and 174 laboratories. These laboratories provided the EQAs with assays of 4,113 blind panels containing 10,538 negative samples (Table 1). The information provided on the laboratories' methods was incomplete, but it appears that 99.8% of the panels were assayed with RT-PCR and 0.2% with RT-Loop-mediated Isothermal Amplification or RT-Reverse Polymerase Amplification. In the EQAs where this information was provided, 17% of the RT-PCR were conventional and 83% were real-time.

In seven EQAs the FPRs were below detection limits and in two we were only able to determine a range. FPRs in each EQA ranged from 0-16.7%; 3.2% of the 10,538 negative samples were reported as positive. There was no correlation between FPR and Year over 2004-2019 ($r=0.147$, $p=0.346$). The median and interquartile range were lower for the full than for the reduced data set, so we used the full data set's 25th percentile (0.8%), median (2.3%) and 75th percentile (4.0%) values as inputs to the analyses (Appendix pp. 3–4).

Published analyses of SARS-CoV-2 testing indicate a range of FNRs from 0% to 40.7% (Appendix pp. 5-6); we used 20% as input to the model and performed a sensitivity analysis over the range of 0-40% (Appendix p. 7). Data on the cumulative numbers of individuals tested and individuals that tested positive in South Korea and New York State through April 22, 2020 were additional inputs. The resulting positivity rates and FDRs are shown in Fig. 1.

For Fig. 1A & 1B, test positivity was calculated as the ratio between the reported numbers of individuals that had tested positive and the number that had been tested. In South Korea, positivity grew from less than 1% before February 20 to a peak of 5.7% on March 2, and then declined to 1.9% by April 22 (Fig. 1A). Assuming an FNR of 20% and FPR of 0.8%, the FDR—the percentage of the positive test results expected to be false positives—ranges from 13.5% on March 2 to 42% on April 22. With FPRs of 2.3% and 4.0% the FDR is higher, reaching 100%—meaning that the expected rate of false positives is equal to or greater than the actual number of positives recorded—on and after April 6 and March 10 respectively (Fig. 1C). In New York State, test positivity grew from 13% on March 13 to 41% on April 8, then declined slightly (Fig. 1B). With FNR=20% and FPR=8%, the FDR ranges between 1% and 5% over this period; with FPR=2.3% the range is 3-15%, and at FPR=4.0% the range is 6-27% (Fig. 1D).

In most countries testing programs targeted individuals that had traveled to high-risk areas, were exposed to individuals that had tested positive, or showed symptoms consistent with COVID-19, and this bias made these data unsuitable for estimating the

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

prevalence of SARS-CoV-2 in the general population. However, one study used data from a quasi-random sampling of the Icelandic population to estimate population prevalence: by dividing the 100 individuals that tested positive by the 13,080 individuals that were tested, the study estimated that 0.8% of the general population was infected.[3] Since FPRs and FNRs were not included in the analysis, we recalculated with an FPR of 0.8% and an FNR between 0 and 40%. The expected number of false positives is 105, slightly greater than the 100 positive results actually reported, so the corresponding estimate of prevalence in the general population would be zero. Varying the FNR over the indicated range has negligible effect.

A small, "accidental" quasi-randomized study was conducted when physicians at a New York City hospital decided to test for SARS-CoV-2 in all obstetrical patients admitted for delivery.[4] They tested 214 patients admitted between March 22 and April 4, 2020, 33 of whom tested positive; 207 of the tested patients and 26 of those that tested positive showed no symptoms of COVID-19 prior to discharge, with a median stay of 2 days. The study reported that more than one out of eight asymptomatic patients (26/207) admitted for delivery were positive for SARS-CoV-2, that these results might be generalizable to regions with similar rates of infection if not to others, and that the true prevalence of infection might be higher because of false negative test results. Recalculating with the test positivity rate of 33/214=15.4%, FNR=20%, and FPRs=0.8%, 2.3%, and 4.0%, and assuming that the seven symptomatic patients were true positives, then the ratio of infected asymptomatic patients to all asymptomatic patients is estimated as 1 to 8.4, 1 to 9.4, and 1 to 11.1, respectively. Note that this calculation takes account of false negatives; varying the FNR from 0 to 40% has little effect on these estimates.

We found two studies that analyzed the SARS-CoV-2 asymptomatic ratio. The first reported on 565 Japanese citizens evacuated from Wuhan, China in February 2020.[5] On arrival in Japan all were tested for SARS-CoV-2 by RT-PCR and 13 tested positive, of which 4 were asymptomatic with no symptoms observed up to 30 days after departure from Wuhan. The study calculated the asymptomatic ratio as 4/13=30.8%, suggesting that this was indicative of the general asymptomatic ratio for COVID-19. We recalculated with the test positivity rate of 13/565=2.3, FNR=0-40%, and FPR=0.8%, yielding an expected number of false positives is between 4.4 and 4.5, greater than the number of asymptomatic individuals that tested positive. If the symptomatic evacuees were true positives, then the estimated asymptomatic rate is zero. With FPR=2.3% the expected number of false positives is 13, equal to the number that tested positive, and the estimated prevalence of SARS-CoV-2 infection among the evacuees is zero.

The second study reported on 3,063 passengers and crew of the *Diamond Princess* cruise ship who were tested for SARS-CoV-2 by RT-PCR;[6] 634 tested positive, including 320 who were asymptomatic at the time of testing. The study used a statistical model to distinguish presymptomatic from asymptomatic members of this group and concluded that 35% of the 320 were true asymptomatics, yielding an asymptomatic ratio of 17.9%. Assuming that the expected false positives are to be found in this asymptomatic group, with a test positivity rate of 634/3,063=20.7%, FNR=0- 40%, and

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

FPR=0.8%, the expected number of false positives is between 18 and 20 and the estimated asymptomatic ratio is 14.8-15.3%; at FPR=2.3% the estimated ratio is 8.9-10.3%, and at FPR=4.0% it is 1.9-4.3%.

Discussion

Though several studies addressed FNRs in SARS-CoV-2 RT-PCR testing, we found little discussion of FPRs. PCR amplification of nucleic acids contributes to test sensitivity but also creates vulnerabilities to minute levels of sample contamination, which can produce false positives that are indistinguishable from true positives. Our meta-analysis of EQAs of similar diagnostic tests found FPRs with an interquartile range of 0.8-4.0%. These false positives were probably not generated by cross-reactivity, since test protocols are typically tested against the likeliest reactants including similar viruses, and because many tests target multiple genomic regions. Nor were they likely to be due to reagent contamination during manufacture, which in most cases would be detected by negative controls.

Rather, the likeliest source of these false positives is sample contamination or human error. Samples can be contaminated by a positive sample analyzed at the same time (cross-contamination), or more likely by target genes amplified from prior positive samples or positive controls (carryover contamination). False positives can also be produced by sample mix-ups[7] or data entry errors.

We investigated the effect of a range of FPRs from 0.8% to 4.0%, corresponding to the interquartile range for the EQAs in our meta-analysis. As an estimate of the possible range of effects this FPR range may be conservative because (1) higher values were reported by several EQAs; (2) the EQAs assessed only intra-laboratory error; any false positives generated by contamination or human error during sampling[8-10] or sample transport would be additive to the FPRs estimated from the EQAs; and (3) the emergency implementation of SARS-CoV-2 testing—including rapid expansion of sampling capacity employing some novel sampling procedures and newly-trained staff, scarcities of personal protective equipment, novel diagnostic assays and controls, and laboratories pressed to operate beyond their normal capacities—suggest a greater potential for false positives than during normal operations.[8,11] Contamination of test kits produced by the U.S. Center for Disease Control and Prevention (US CDC),[12] and reported cross-reactivity with primer-probe sets used in some SARS-CoV-2 assays[13] may be suggestive of the consequences of a rushed response.

The FPR is the proportion of uninfected individuals that test positive. Thus even a small FPR can generate a significant number of false positives if uninfected individuals are a large fraction of the test population, as is generally the case when test positivity is low. This can be seen in our re-analyses of four studies. In the Iceland and Japanese evacuees studies, which had test positivity rates of 0.8% and 2.3% respectively, our re-calculations using even a low FPR greatly changed the results, reducing the estimates of population prevalence and asymptomatic ratios to zero. In contrast, in studies of women admitted for delivery and the *Diamond Princess* cruise ship, which had test

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

positivity rates of 15.4% and 20.7% respectively, re-calculations using the same FPR reduced the studies' results by only 5% and 16% respectively.

Our recalculations were based on estimated FPRs and FNRs, which even if accurate with regard to the mean may be higher or lower in some localities than in others. For example in Iceland, although our recalculations indicate zero or near-zero prevalence in the general population, certain aspects of the population data—that a larger fraction of the group that tested positive had traveled to other countries, had contact with an infected individual, and had reported some symptoms; and that no children tested positive—suggest that many of the individuals that tested positive were actually infected.

Our analysis of two regional data sets demonstrates the potential for a high frequency of false positives, especially when test positivity is low. In New York State, with FNR=20% and FPR=0.8%-4.0%, estimated FDR declined from 5-17% when positivity was lower in mid-March to 1-5% at the positivity peak on April 11 (Fig. 1D). Test positivity was lower in South Korea. From the date of peak positivity to April 12 the estimated FDR rose from 14% to 37% if FPR=0.8%; it rose from 40-71% to 100% by different dates if FPR=2.3% or 4.0%, when the expected number of false positives was high enough to account for all of the positive test results (Fig. 1C). If included in calculations, a high frequency of false positives reduces estimates of prevalence, which affect other key statistics, for example by increasing estimates of the hospitalization rate and death rate,

Thus FDR rises as test programs sample populations with fewer infected individuals. This can result from programs continuing to implement broad-scale testing even as the prevalence of the virus in the general population declines, as appears to be the case in South Korea. Or it can result from expanding the scale of testing to include individuals who are less likely to be infected. Calls for mass testing for SARS-CoV-2 should be evaluated in light of the potential for the positive test results from broad-scale testing to be substantially comprised of false positives. Though broadening the scale of testing does generate more data, the quality of the generated data declines as it becomes increasing contaminated by false positives.

At the case level, the FDR is the probability that a positive result is incorrect. The fact that substantial FDRs result from applying modest FPRs to various real-world test data has implications for case management. False-positive test results may cause unnecessary stress and result in unnecessary, weeks-long isolation of individuals, including isolation of health-care and other critical workers. Human and other resources are wasted if used to track and test close contacts of false positives. In hospitals, long-term care facilities, and prisons, false-positive patients, residents, and inmates may be moved into isolation wards or living areas along with infected individuals, placing them at greater risk of infection.[14] Similarly, in some regions, false-positive individuals living at home may be pressured or required to move into isolation facilities for people infected with SARS-CoV-2.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

In hospitals, there can be wasteful consumption of personal protective equipment in caring for false positive patients, while at the same time the unnecessary use of such equipment and other protective measures may impede the care of the patient. In respiratory patients, a false positive test result may impede a correct diagnosis, delaying or depriving such patients of appropriate treatment. The presence of false-positive patients may hinder the development of improved medical care for COVID-19 patients based on clinical experience, because incorrectly diagnosed patients introduce noise into clinical observations. Over the longer term, if antibody or antiviral treatments become available for COVID-19 patients, or if prophylactic treatments are developed for asymptomatic or mildly symptomatic individuals that have tested positive, false-positive individuals may be subjected to medically inappropriate treatments.[15] Clinical trials of potential treatments could lose statistical power by unwittingly enrolling false-positive individuals, who would be exposed to potentially harmful side effects without any mitigating potential for benefit.

False positives affect the interpretation of test results. The Patient's Fact Sheet for the US CDC's SARS-CoV-2 test states that there is only "a very small chance" that positive tests result could be wrong, but suggests that negative results could well be wrong.[16] Public health officials have similarly suggested that positive test results are absolutely reliable but negative results are untrustworthy.[17] In reality the opposite is true over a wide range of likely scenarios. For example, antibody tests indicate that the prevalence of SARS-CoV-2 infections in residents of Santa Clara County, California is 2.49% to 4.16%.[18] At those prevalences, FPR=0.8%, and FNR=40%, the probability that a positive result is false is 24-34% while the probability that a negative result is false is 1.0-1.7%: it is 14-34 times more likely that a positive than a negative result is wrong. Positive results are more likely to be wrong than negative results at prevalences from 0-15%; with higher FPRs or lower FNRs, this remains true at higher prevalences. The explanation for this is straightforward. The FPR acts on samples from the uninfected fraction of the population, returning positive results for some uninfected individuals, while the FNR acts on the infected fraction, returning negative results for some infected individuals. When prevalence is low, the uninfected fraction is much greater than the infected fraction, so that even a low FPR can have a considerably larger effect than a high FNR.

We suspect that under-appreciation of the potential impact of false positives in SARS-CoV-2 testing has two sources. First, because PCR-based diagnostic protocols are effectively designed to eliminate false positives due to cross-reactivity, some individuals assume that false positives have been eliminated from the test process. In practice, however, minute levels of contamination, which are extremely challenging to control, can produce false positives in PCR-based tests, and the potential for human error in sample handling or records management can probably never be entirely eliminated. EQAs regularly found significant FPRs, from whatever cause, in PCR-based assays. Second, there appears to be little understanding of the relationship between FPR and FDR, in which a small FPR can produce a large FDR if the test positivity rate is low. This was evident when US health officials described a study as reporting that a SARS-CoV-2 test had a 47% FPR, when it actually reported a 47% FDR.[19]

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

A limitation of our study was the application of FPR estimates derived from EQAs of tests for other RNA viruses to SARS-CoV-2 test data. However, there is no apparent reason that SARS-CoV-2 protocols should produce significantly different FPRs, and we derived our model inputs conservatively from the EQA data. The EQA yielded FPR estimates on a per sample basis, and in analyzing the regional data we applied these to data aggregated on an individual basis, with some individuals tested multiple times. The individual-basis FPR may differ from the sample-basis FPR, but if the number of repeat tests is small relative to total tests the difference should be small. We used uncertain FNR estimates from the literature, but showed that varying FNR has little effect on our analyses.

There is much that we could do to mitigate the impact of false positives. First in importance is improved awareness. With understanding of the potential frequency of false positives and their effects, clinicians, epidemiologists and those involved in critical public health decisions can avoid many of the pitfalls described in this paper. Second, EQAs of SARS-CoV-2 diagnostic laboratories could be implemented to improve FPR estimates, and if designed appropriately, provide insight into the causes of intra-laboratory false positives and facilitate their reduction. Efforts could also be made to assess the rate of false positives in sampling and sample transport. Third, false positives could be greatly reduced—though at the cost of increasing false negatives and consuming testing resources—by requiring two positive tests to count an individual as positive. This would ideally be done by taking two samples or splitting a sample at the point where it is taken, and sending each to different laboratories for analysis. Alternatively, false positives due to contamination at low concentrations could potentially be addressed by lowering the maximum Ct value for positive results, or by using primer-probe sets that are less sensitive. The resulting increase in FNR might be a reasonable trade-off for reduced FPR where test prevalence is low. Fourth, the rate of false positives in SARS-CoV-2 RT-PCR tests can be assessed retrospectively by antibody tests, though the accuracy of those tests may be similarly affected by false positives.

Contributors

AC designed and conducted the analyses. AC and BK collaborated on the review of clinical implications and writing the report.

Funding and Role of Funding Organization

There was no external funding for this research.

Declaration of interests

We declare no competing interests.

Acknowlegments

We thank Michael Milgroom for his help with statistical analysis and writing the report, Elisa Liberti for help with data acquisition, and Dominic Chow for reviewing the manuscript.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

References

1 Korea Centers for Disease Control and Prevention. Updates on COVID-19. https://www.cdc.go.kr/board/board.es?mid=&bid=0030 (accessed Apr 22, 2020).

2 The COVID Tracking Project. New York/Historical Data for New York. https://covidtracking.com/data/state/new-york#historical (accessed Apr 22, 2020).

3 Gudbjartsson DF, Helgason A, Jonsson H, et al. Spread of SARS-CoV-2 in the Icelandic Population. *New Engl J Med* 2020; published online Apr 14, 2020. DOI: 10.1056/NEJMoa2006100.

4 Sutton D, Fuchs K, D'Alton M, Goffman D. Universal screening for SARS-CoV-2 in women admitted for delivery. *New Engl J Med* 2020; published online Apr 13, 2020. DOI: 10.1056/NEJMc2009316.

5 Nishiura H, Kobayashi T, Suzuki A, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int J Infect Dis* 2020; published online Mar 13, 2020. DOI: 10.1016/j.ijid.2020.03.020.

6 Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the *Diamond Princess* cruise ship, Yokohama, Japan, 2020. *Euro Surveill* 2020; **25**(10). DOI: 10.2807/1560-7917.ES.2020.25.10.2000180.

7 Lau KA, Theis T, Gray J, Rawlinson WD. Ebola preparedness: diagnosis improvement using rapid approaches for proficiency testing. *J Clin Microbiol* 2017; **55**: 783–90.

8 Qian Y, Zeng T, Wang H, et al. Safety management of nasopharyngeal specimen collection from suspected cases of coronavirus disease 2019. *Int J Nurs Sci* 2020; published online Apr 4, 2020. DOI: 10.1016/j.ijnss.2020.03.012.

9 Wang M, Wu Q, Xu W, et al. Clinical diagnosis of 8274 samples with 2019-novel coronavirus in Wuhan. *medRxiv* 2020. https://doi.org/10.1101/2020.02.12.20022327 (accessed Apr 24, 2020).

10 Carbone M, Green JB, Bucci EM, Lednicky JA. Coronaviruses: facts, myths, and hypotheses. *J Thorac Oncol* 2020; **15**: 675–8.

11 Lippi G, Simundic AM, Plebani M. Potential preanalytical and analytical vulnerabilities in the laboratory diagnosis of coronavirus disease 2019 (COVID-19). *Clin Chem Lab Med* 2020; published online Mar 16, 2020. DOI: 10.1515/cclm-2020-0285.

12 Willman D. Contamination at CDC lab delayed rollout of coronavirus tests. *Washington Post*, Apr 28, 2020. https://www.washingtonpost.com/investigations/contamination-at-cdc-lab-delayed-rollout-of-coronavirus-tests/2020/04/18/fd7d3824-7139-11ea-aa80-c2470c6b2034_story.html (accessed Apr 24, 2020).

13 Vogels CBF, Brito AF, Wyllie AL, et al. Analytical sensitivity and efficiency comparisons of SARS-COV-2 qRT-PCR assays. *medRxiv* 2020. https://www.medrxiv.org/content/10.1101/2020.03.30.20048108v2.full.pdf+html (accessed Apr 24, 2020).

14 Wang Y, Wang Y, Chen Y, Qin Q. Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. *J Med Virol* 2020; 92(6): 568–76.

15 Molloy PJ, Persing DH, Berardi VP. False-positive results of PCR testing for Lyme disease. *Clin Infect Dis* 2001; **33**: 412–3.

16 U.S. Centers for Disease Control and Prevention. Fact Sheet for Patients: CDC - 2019-nCoV Real-Time RT-PCR Diagnostic Panel. https://www.cdc.gov/coronavirus/2019-ncov/downloads/Factsheet-for-Patients-2019-nCoV.pdf (accessed Apr 24, 2020).

17 Johnson CY. A 'negative' coronavirus test result doesn't always mean that you aren't infected. *Washington Post*, Mar 26, 2020. https://www.washingtonpost.com/science/2020/03/26/negative-coronavirus-test-result-doesnt-always-mean-you-arent-infected/ (accessed Apr 24, 2020).

18 Bendavid E, Mulaney B, Sood N, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. *medRxiv* 2020. https://doi.org/10.1101/2020.04.14.20062463 (accessed Apr 24, 2020),

19 Harris R. In defense of coronavirus testing strategy, Administration cited retracted study. *National Public Radio (NPR)*, Mar 26, 2020. https://www.npr.org/sections/health-shots/2020/03/26/822084429/in-defense-of-coronavirus-testing-strategy-administration-cited-retracted-study (accessed Apr 24, 2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

Table 1. False positive rates in external quality assessments (EQAs) of diagnostic laboratories assaying RNA viruses.

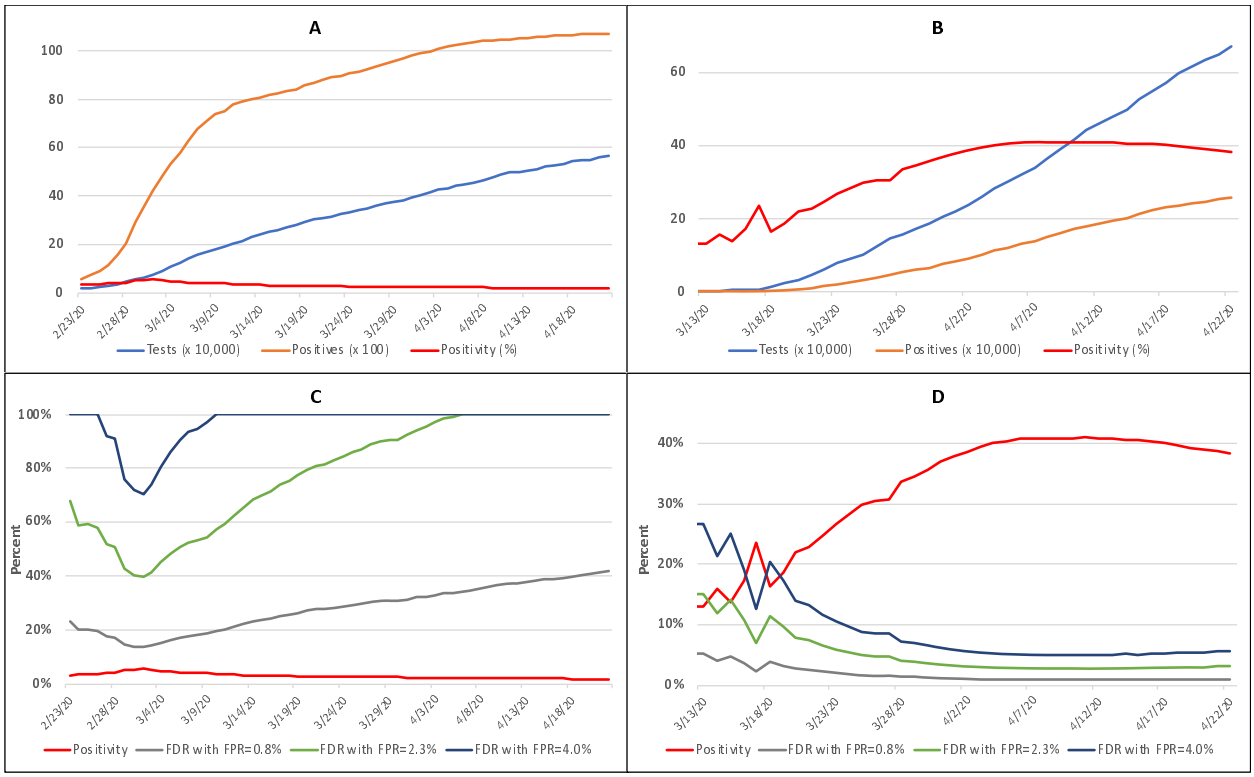
Study	Viral Target	Date of EQA	Labor-atories	Negative Samples	False Positive Rate
Drosten *et al.* 2004	SARS-CoV	2004?	58	174	2.3-6.9%
Pas *et al.* 2015	MERS-CoV	2014	99	1,134	1.0%
Zhang *et al.* 2016	MERS-CoV	2015?	56	168	<0.6%[e]
Zhou & Luo 2018	MERS-CoV	2017?	49	49	<2.0%[e]
WHO 2008-2020[a,d]	Influenza A viruses	2007-2019	64-174	114-332	<0.6[e]-7.0%
Caliendo *et al*. 2006	Hepatitis C virus	2005?	5	119	3.4%
Laperche *et al.* 2007	Hepatitis C virus	2006?	20	21	<4.8%[e]
Tvorogova *et al.* 2009[b,d]	Hepatitis C virus	2005-07	78-104	534-728	2.1-7.0%
LeGal *et al.* 2016	Hepatitis Delta virus	2015?	28	112	5.4%
Panning *et al.* 2009	Chikungunya virus	2007?	31	108	1.9-5.6%
Jacobsen *et al*. 2016	Chikungunya virus	Sep 2014	56	297	8.1%
Soh *et al.* 2015	Chikungunya, Dengue	Feb-May 2015	20	40	2.5%
Pok *et al.* 2015	Dengue virus	May-Jul 2013	16	16	6.3%
Charrel *et al.* 2017	Zika virus	Oct 2016	50	504	2.8%
Escadafal *et al.* 2013	Rift Valley Fever virus	2012	30	117	3.4%
Zhang *et al.* 2015	Measles virus	Aug 2014	41	123	0.8%
Wang *et al.* 2015	Ebola virus	Dec 2014	19	60	<1.7%[e]
Ellerbrok *et al*. 2017	Ebola virus	Aug 2014	82	317	0.3%
Lau *et al.* 2017[c,d]	Ebola virus	2014-16	3-9	3-9	<11.1[e]-16.7%
Reusken *et al.* 2019	4 arboviruses	Nov 2017	51	204	4.9%

Laboratories = the number of laboratories participating in the EQA; Negative Samples = the total number of negative samples in the blind panels analyzed by those laboratories. Citations are listed in the Appendix pp. 8–9.

[a] Comprises 17 EQAs.
[b] Comprises 6 EQAs.
[c] Comprises 3 EQAs.
[d] Shows the range in number of laboratories, negative samples and false positive rates per EQA.
[e] Below the indicated detection limit; treated as zero in the analyses.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.26.20080911; this version posted May 1, 2020. 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 .

Figure 1. Cumulative numbers of individuals tested (Tests), individuals that tested positive (Positives), and test positivity (Positivity), for A) South Korea and B) New York State; and the resulting false discovery rate (FDR) based a 20% false negative rate and the indicated false positive rates (FPR), for C) South Korea and D) New York State.



New York State counts an individual tested on n different days as n individuals. Note the different scales for Tests and Positives in Panel A.