medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .

Reconstructed diagnostic sensitivity and specificity of the RT-PCR test for COVID-19

Nikhil S. Padhye[1]

[1] The University of Texas Health Science Center at Houston, Houston, Texas, USA

Corresponding Author:
Nikhil Padhye[1]
6901 Bertner Avenue, Suite 560, Houston, TX, 77345, USA
Email address: Nikhil.S.Padhye@uth.tmc.edu

Abstract

Background. Real-time reverse transcription polymerase chain reaction (RT-PCR) targeting select genes of the SARS-CoV-2 RNA has been the main diagnostic tool in the global response to the COVID-19 pandemic. However, the diagnostic accuracy of the test has not been studied systematically outside of the laboratory setting. The aim of this study is to provide estimates of the diagnostic sensitivity and specificity of the RT-PCR test developed by China CDC.
Methods. The study design is a secondary analysis of published findings on 1014 patients in Wuhan, China, of whom 601 tested positive and 413 were negative for COVID-19. Sensitivity and specificity were reconstructed using a Bayesian approach from probabilistic knowledge of the diagnostic errors. Predictive values of the test were calculated, resulting in estimates for the number of confirmatory tests that are needed for establishing the presence or absence of COVID-19, depending on the prior probability of a patient having the disease.
Results. The sensitivity of the RT-PCR diagnostic test was estimated to be 0.777 (95% CI: 0.715, 0.849), while the specificity was 0.988 (95% CI: 0.933, 1.000). The confidence intervals include sampling error in addition to the error due to probabilistic knowledge of the data.
Discussion. The Chinese version of the RT-PCR test had a conspicuous rate of false negative results, likely missing between 15% and 29% of patients with COVID-19. For a patient with a prior probability of COVID-19 greater than 18%, at least two negative test results would be needed to lower the chances of COVID-19 below 5%. Caution is advised in generalizing these findings to other versions of the RT-PCR test that are being used in diverse geographic regions.

Introduction

The cause of a disease outbreak that began in Wuhan, China in the last quarter of year 2019 was later identified as a novel coronavirus, labeled SARS-CoV-2 since it can cause severe acute respiratory syndrome. The disease associated with SARS-CoV-2 has been termed COVID-19 (1). The publication of the SARS-CoV-2 genome (2) led to the rapid development in January 2020 of real-time reverse transcription polymerase chain reaction (RT-PCR) tests for the

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/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

diagnosis of COVID-19 while avoiding cross-reactions to other known coronaviruses. One version was developed in China that targeted the *ORF1ab* and *N* genes of viral RNA (3) while another version was developed in Germany that targeted the *RdRp*, *E*, and *N* genes (4). Real-time RT-PCR tests were developed and implemented thereafter by many laboratories around the world (5–7), even as COVID-19 became a global pandemic that continues to spread rapidly at this time. A listing of tests and protocols is being maintained online by the World Health Organization (8).

The rapid development and deployment of RT-PCR tests has been essential for the ability to measure and control the spread of SARS-CoV-2. However, the urgency of the pandemic has not allowed time for reliable and adequately powered clinical studies to be conducted to measure the diagnostic limitations of the RT-PCR test. At present, the bulk of knowledge about the sensitivity and specificity of RT-PCR is based on laboratory measurements that have goals related to the minimum threshold of detection of viral loads and the required number of thermal cycles of the chain reaction (9). Some attention has been given to the viral distribution by physical location, such as the differences in positive rates of RT-PCR in nasopharyngeal versus oropharyngeal swabs, or in the sputum and bronchoalveolar lavage fluid (10,11). Other factors that can impact the diagnostic success of RT-PCR include the timing of the test relative to disease onset, adequacy of the volume of fluids collected in the swab, and deviations from the laboratory-recommended protocol under real-world conditions. In terms of clinical decision-making, any of the causes of failure of the test can lead to incorrect diagnoses due to false positives and false negatives. This issue has received media attention (12) and a recent editorial written by a professor of medicine in an influential US newspaper (13) urged physicians to beware of false negatives of diagnostic testing for COVID-19 while acknowledging that reliable data on rates of false negatives were not yet available.

This study provides a timely assessment of the diagnostic sensitivity and specificity of RT-PCR that is based on a sample of 1014 patients in Wuhan, China (14). The original study had the aim of measuring the accuracy of chest CT imaging for diagnosis of COVID-19 and they assumed that RT-PCR was the gold standard. However, the authors provided additional information about the status of patients that allows for the reconstruction of the sensitivity and specificity of RT-PCR in the context of clinical decision-making. The knowledge of the efficacy and limitations of RT-PCR, even if known for only the Chinese version of the test at this time, can be expected to provide a valuable reference for medical practitioners and researchers at the frontlines of the fight against the pandemic. The findings have implications for policy makers as well, because policies for pandemic control in conditions of limited availability of tests that have a high rate of false negatives can be starkly different from policies in the presence of an abundant supply of a diagnostic test with excellent predictive values.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

Data and Methods

Data

Data from published findings (14) have been used in this study. The study included 1014 patients suspected of having COVID-19 in Wuhan, China, who underwent RT-PCR and chest CT imaging diagnostic tests during a 30-day period in the months of January and February, 2020. The mean age was reported to be 51 ±15 years, and 46% were male. Throat swab samples were collected and the RT-PCR assays were reported to have used TaqMan One-Step RT-PCR kits from Shanghai Huirui Biotechnology Co., Ltd., or Shanghai BioGerm Medical Biotechnology Co., Ltd., both of which were reported to have been approved for use by China Food and Drug Administration. RT-PCR tests were positive for 601 patients (59.3%) and negative for the other 413 patients (40.7%). Although these tests were treated as the gold standard for comparison with chest CT imaging, the study authors provided valuable additional information. Patients who had negative RT-PCR tests but positive tests from chest CT were reassessed on the basis of clinical symptoms, CT features, and serial CT scans. The study staff concluded that among patients with negative RT-PCR tests, 147 could be classified as *highly likely* cases of COVID-19 and another 103 could be classified as *probable* cases of COVID-19. Moreover, among the 601 patients with positive RT-PCR tests, 21 patients were classified negative for COVID-19 from chest CT imaging. In this study, these 21 cases are assumed to have a low chance of being false positives, which is in alignment with the implicit assumption of the original study, indicated by their choice of RT-PCR as the reference for comparison with chest CT imaging. Apart from its role in the identification of probable false positives and negatives of RT-PCR, the chest CT imaging data is ignored in this study. In summary, the data are composed of firm knowledge of RT-PCR test results and probabilistic knowledge of the numbers of false positives and false negatives. The ranges and notations used for the true and false positives and negatives are presented in Table 1.

	COVID-19 present	COVID-19 absent	Row Sums
Positive RT-PCR	$N_1 - n_1$	n_1	$N_1 = 601$
Negative RT-PCR	n_2	$N_2 - n_2$	$N_2 = 413$
Column Sums	$N_1 - n_1 + n_2$	$N_2 - n_2 + n_1$	$N_1 + N_2 = 1014$

Table 1. Counts and notation for the joint distribution of RT-PCR test result and presence or absence of COVID-19. Maximum range of variation for n_1 is [0, 21] and the maximum range for n_2 is [0, 250], but probable ranges are much narrower. The 95% CI for n_1 was found to be [0, 9] and for n_2 it was [136, 197].

Statistical Analysis

A modified Bayesian approach was adopted to estimate the uncertainty that arose from imprecise knowledge of the data. The data, denoted X, consisted of the 2x2 contingency table that represented the observed joint distribution of the RT-PCR decision (positive or negative) and the binary disease status (COVID-19 present or absent). Since the number of positive and negative tests were known, the data were uniquely defined from the number of false positives, n_1, and

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

number of false negatives, n_2, i.e. $X = X(n_1, n_2)$. The uncertainty in the data stemmed from the uncertainty in values of the random variables, n_1 and n_2. The distributions of n_1 and n_2 were estimated from the level of confidence expressed about false identifications. This procedure has similarities to fuzzy logic in which linguistic uncertainties about terms such as *highly likely* and *probable* are represented by membership functions (15). Although it is motivated by fuzzy logic, the treatment used here is strictly based in probability theory. The starting point was an informative distribution defined on the probability space of a diagnostic error. Upon calculating the resulting distributions of n_1 and n_2, the estimated values and distributions of sensitivity and specificity were derived from the joint distribution of n_1 and n_2. More details are given below.

The approach is thus Bayesian in the computational sense; it starts with an informative distribution, akin to an informative prior distribution, and ends with a distribution of the desired parameters (16). However, the terminal points of the analysis do not describe the distribution of the same parameters, so use of the terms *prior* and *posterior* distributions has been avoided. Moreover, the likelihood function of the data that is calculated here has a different interpretation than the one obtained in normal conditions when the data are firmly known. For any given probability of false identification, the likelihood of the number of false positives or negatives is obtained from the binomial distribution. The number of trials for the binomial distribution are known from the data, while the binomial probability parameter, denoted here by θ, arises from a distribution that represents the degree of confidence expressed about the false identifications. Thus, the distribution of n_1 is expressed by:

$$\Pr(n_1) = \sum_{\theta_L} \Pr(n_1 | \theta_L) \Pr(\theta_L),$$

where θ_L is the chance of a case being false positive, the subscript L refers to a *low* chance, $\Pr(\theta_L)$ is the distribution of that chance over the probability space [0, 1], and $\Pr(n_1 | \theta_L)$ is the binomial distribution with probability θ_L and number of trials given by the maximum range for n_1, which is known to be 21. Similarly, the distribution of n_2 is expressed by:

$$\Pr(n_2) = \sum_{\theta_H, \theta_M} \Pr(n_H | \theta_H) \Pr(\theta_H) \Pr(n_M | \theta_M) \Pr(\theta_M),$$

where θ_H and θ_M are the *high* and *medium* chances of a case being *highly likely* to be a false negative and *probable* false negative, respectively. The sum of n_H and n_M equals n_2 and the distributions of n_H and n_M are binomial, given values of θ_H and θ_M along with the number of binomial trials, which are known to be 147 and 103, respectively. Beta distributions, which often serve as conjugate distributions for the binomial distribution, were used for θ_L, θ_M, and θ_H to describe low, medium, and high levels of confidence in the diagnostic errors that were identified (17).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

The data $X(n_1, n_2)$ were uniquely specified by n_1 and n_2, as were the sensitivity, S_1, and specificity, S_2. In particular, $S_1(n_1, n_2) = (N_1 - n_1) / (N_1 - n_1 + n_2)$ and $S_2(n_1, n_2) = (N_2 - n_2) / (N_2 - n_2 + n_1)$, where the values of N_1 and N_2 are known. Therefore, the joint distribution $\Pr(n_1, n_2)$ provided a mapping to the distributions of the sensitivity and specificity. The careful considerations described up to this point resulted in estimates of the expected values of sensitivity, specificity and a measure of the uncertainty in their values that arose from imperfect knowledge of the data. Another source of uncertainty is due to sampling error, which was estimated using established methods for the standard error for proportions.

Lastly, this study evaluated the predictive values of the test that provide the chance of disease in a patient conditional upon results of the diagnostic test. Sensitivity is a conditional probability that can be reversed using the Bayes formula to provide the positive predictive value of the test:

$$\Pr(\text{COVID-19}|\text{Positive RT-PCR}) = \frac{S_1 \Pr\,(\text{COVID-19})}{S_1 \Pr(\text{COVID-19}) + (1 - S_2)(1 - \Pr(\text{COVID-19}))}.$$

A similar formula can be expressed for the probability of COVID-19 in a patient even if the test is negative, i.e. Pr (COVID-19 | Negative RT-PCR). In the above equation, Pr(COVID-19) on the right hand side can be interpreted as the prevalence of the disease when the testing is being done in the general population. More generally, Pr(COVID-19) is the prior probability of presence of the disease. This is also the more appropriate interpretation in the present circumstances in which the testing is reserved largely for symptomatic patients. The medical

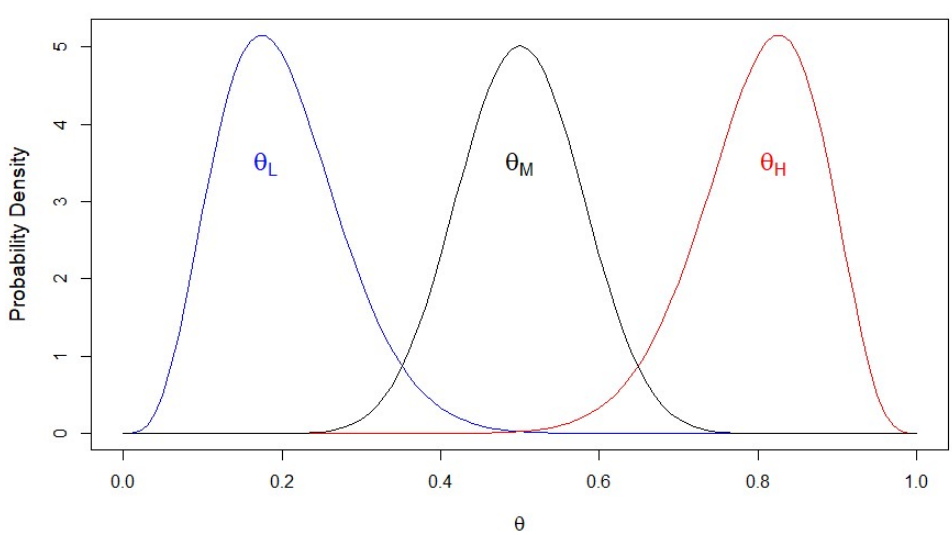

Figure 1. Distributions corresponding to low, medium, and high confidence of identification of diagnostic errors. The beta distribution for θ_L has parameters a=5, b=20. The beta distribution for θ_M has parameters a=20, b=20. The beta distribution for θ_H has parameters a=20, b=5.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

professional may suspect that a patient has COVID-19, which could be quantified into the prior probability. The RT-PCR test is then carried out and the test result decides the posterior probability. Statistical analysis was done using the R programming language (18) in the RStudio software environment (19).

Results

Shape parameters of the beta distributions for low, medium, and high confidence levels for being a false test result were a=5, b=20 for Pr(θ_L), a=20, b=20 for Pr(θ_M), and a=20, b=5 for Pr(θ_H). The median value of θ_L was 0.192, with 10th and 90th percentiles given by 0.105 and 0.306, respectively. The median value of θ_M was 0.500, with 10th and 90th percentiles equal to 0.399 and 0.601, respectively. For θ_H, the median value was 0.808, with 10th and 90th percentiles given by 0.694 and 0.895, respectively. The distributions for θ_L, θ_M, and θ_H are shown in Figure 1.

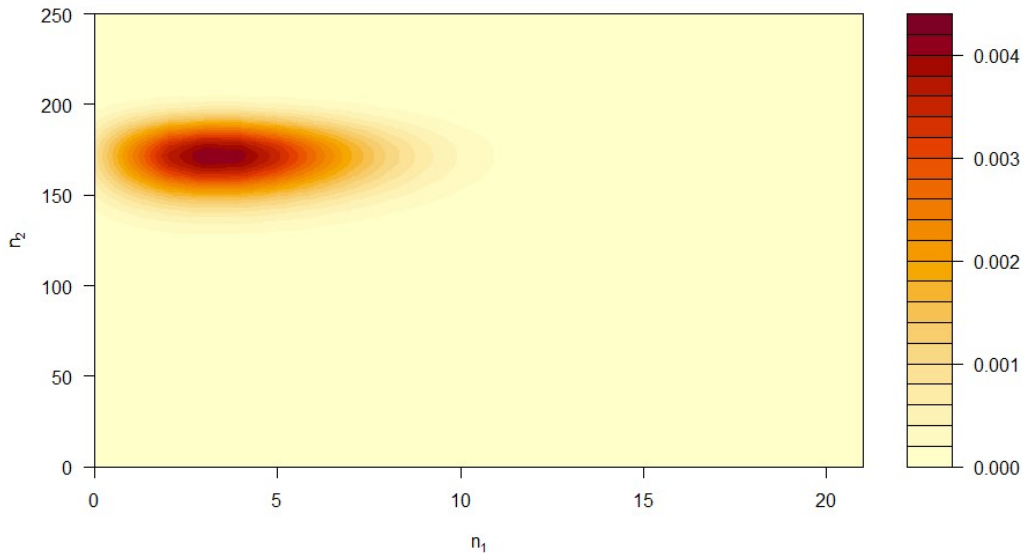

Figure 2. Joint distribution of false positives, n_1, and false negatives, n_2. The color scale refers to the probability of observing any pair of values of n_1 and n_2.

The maximum likelihood of the joint distribution of n_1 and n_2 (Figure 2) was located at n_1 = 3 and n_2 = 172. This solution corresponds to an estimate of 770 patients with COVID-19 and 244 without the disease. Due to the low number of false positives, the expectation value of specificity was high: $\hat{S}_2$ = 0.988. In contrast, the high false negative count was reflected in the lower expectation value of sensitivity: $\hat{S}_1$ = 0.777. The 95% confidence intervals for Pr(n_1, n_2) led to the corresponding limits for Pr(S_1 | n_1, n_2) and Pr(S_2 | n_1, n_2). Sensitivity had a 95% confidence interval from 0.746 to 0.821, while the 95% confidence interval for specificity ranged between 0.958 and 1.000. Figure 3 depicts the distributions of sensitivity and specificity.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

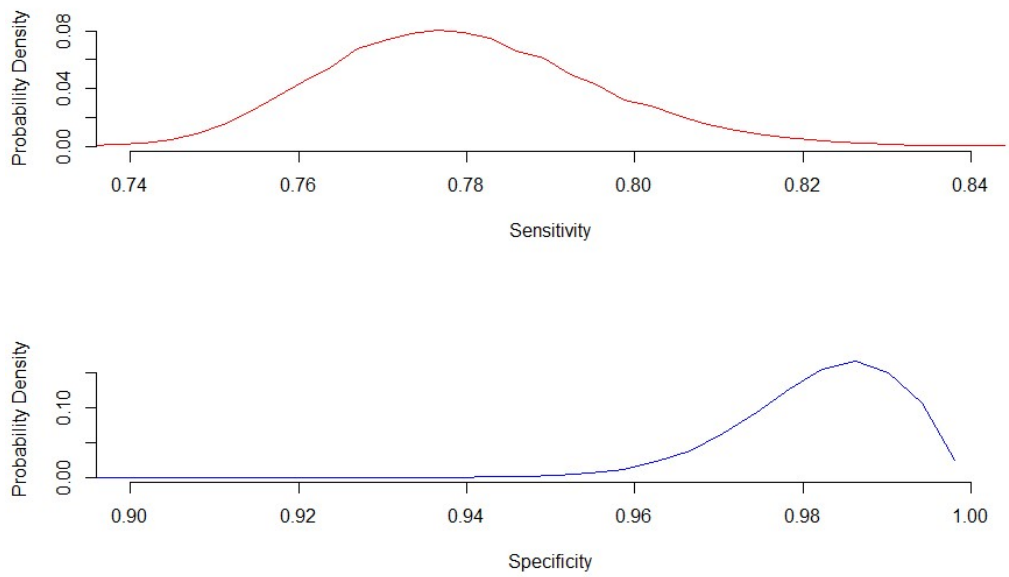

Figure 3. Distributions of sensitivity, $\Pr(S_1 \mid n_1, n_2)$, and specificity $\Pr(S_2 \mid n_1, n_2)$ estimated by the modified Bayesian method described in the study.

The 95% confidence intervals mentioned above provided a measure of the uncertainty that arose from imperfect knowledge of the data. Additionally, the sampling error was estimated from the standard error for proportions evaluated at each of the endpoints of the 95% confidence interval for the data-related uncertainty. The overall 95% confidence intervals that incorporate the two sources of error are shown in Table 2.

		Value	95% CI (data uncertainty)		95% CI (incl. sampling error)	
		(est.)	Lower	Upper	Lower	Upper
Set 1	Sensitivity	0.777	0.746	0.821	0.715	0.849
	Specificity	0.988	0.958	1.000	0.933	1.000
Set 2	Sensitivity	0.778	0.757	0.805	0.727	0.833
	Specificity	0.984	0.963	1.000	0.940	1.000
Set 3	Sensitivity	0.779	0.740	0.842	0.709	0.869
	Specificity	0.983	0.950	1.000	0.923	1.000

Table 2. Estimated values of sensitivity and specificity of RT-PCR for diagnosis of COVID-19. Set 1 refers to the primary choice of beta distributions for low, medium, and high confidence levels of the identified diagnostic errors, while Set 2 and Set 3 are explorations resulting from the assumption of narrower and wider beta distributions. Among columns, the first 95% CI represents the uncertainty arising from imperfect data. The second 95% CI includes sampling error in addition to the data uncertainty.

The impact was explored of selecting different shape parameters for the beta distributions that describe low, medium, and high confidences. First, narrower beta distributions were defined by using shape parameters a=20, b=80 for $\Pr(\theta_L)$, a=80, b=80 for $\Pr(\theta_M)$, and a=80, b=20 for

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

Pr(θ_H). The median values of θ_L, θ_M, and θ_H were 0.198, 0.500, and 0.802, respectively. The span between the 10th and 90th percentiles was approximately 0.1, which may be compared to 0.2 for the distributions described previously. Second, wider beta distributions were defined by using shape parameters a=3, b=10 for Pr(θ_L), a=10, b=10 for Pr(θ_M), and a=10, b=3 for Pr(θ_H). The median values of θ_L, θ_M, and θ_H were 0.217, 0.500, and 0.783, respectively. The span between the 10th and 90th percentiles was approximately 0.29. The estimated values of sensitivity, specificity, and their 95% confidence intervals are shown in Table 2. Point estimates of the parameters showed very little variation, but the choice of narrower/wider beta distributions resulted in somewhat narrower/wider confidence intervals.

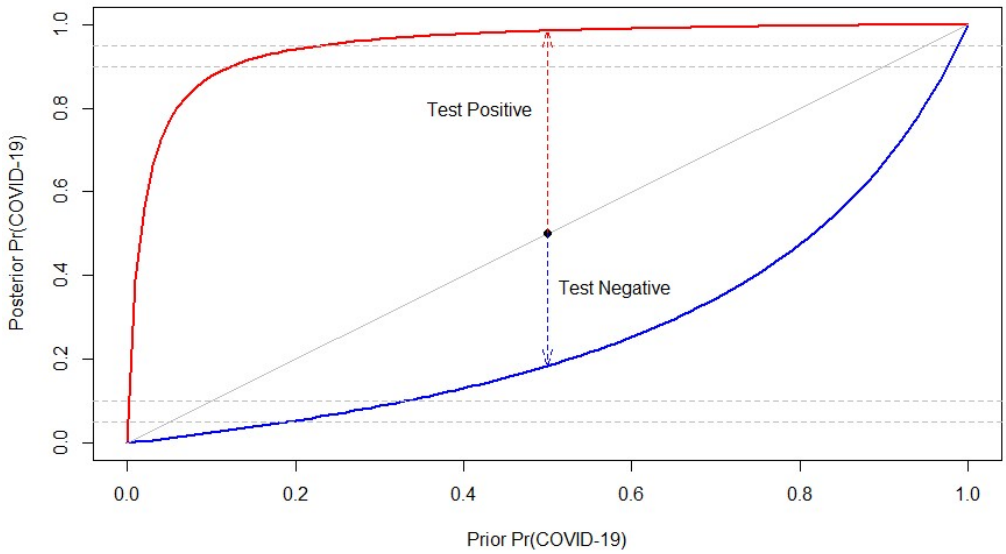

Figure 4. Probability of having COVID-19 after RT-PCR test as a function of the prior probability, or suspected chance that a patient may have the disease. The upper (red) curve applies for a person who tests positive, while the lower (blue) curve applies when the test is negative. The diagonal is the line of equality of prior and posterior probabilities; any point on it may be considered the predicted probability before the diagnostic test is done. The arrows show an example of the changes induced by test results: the prior value of 0.50 jumps to the posterior value of 0.98 after a positive test and drops to 0.18 after a negative test result. The horizontal dashed lines mark 90% and 95% confidences at the top and bottom.

The predictive values of the RT-PCR diagnostic test are shown in Figure 4 for prior probabilities ranging from 0 to 1. The two curves in the figure show the posterior probabilities of the presence of COVID-19 when test results are positive or negative. Additionally, Table 3 displays the number of confirmatory tests that are needed to establish presence or absence of COVID-19 at confidence levels of 90% and 95%. For example, if the prior probability of presence of COVID-19 in a patient is judged to be 0.6, a single negative RT-PCR test would reduce that probability to 0.253. A second negative test would reduce it further to 0.070, which would be sufficient if at least 90% confidence is required to establish absence of the disease. However, for the confidence

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

level of 95%, a third negative test would be needed to lower the probability of COVID-19 below 0.05.

	Posterior Probability of COVID-19		Num. Tests to Establish Presence of COVID-19		Num. Tests to Establish Absence of COVID-19	
Prior Probability of COVID-19	Positive Test	Negative Test	P > 0.90	P > 0.95	P < 0.10	P < 0.05
0.1	0.875	0.025	2	2	1	1
0.2	0.940	0.054	1	2	1	2
0.3	0.964	0.088	1	1	1	2
0.4	0.977	0.131	1	1	2	2
0.5	0.984	0.184	1	1	2	2
0.6	0.990	0.253	1	1	2	3
0.7	0.993	0.345	1	1	3	3
0.8	0.996	0.475	1	1	3	3
0.9	0.998	0.671	1	1	3	4

Table 3. Predictive values of the RT-PCR diagnostic test for COVID-19. Posterior probabilities are shown for prior probabilities ranging from 0.1 to 0.9, along with the number of confirmatory tests needed to establish presence or absence of COVID-19 with confidence levels of 90% and 95%.

Discussion

RT-PCR tests are commonly used for the diagnosis of many influenza viruses and coronaviruses, including the viruses responsible for the 2002-04 SARS coronavirus outbreak, the 2009 H1N1 influenza pandemic, and the 2012 MERS coronavirus outbreak. RT-PCR tests are often treated as the gold standard in comparisons of diagnostic methods, which has led to few sources of reliable data about their diagnostic accuracy in clinical practice. The virus culture process is considered a better standard, but it takes several days in comparison to the few hours needed for RT-PCR tests. In one such comparison (20), RT-PCR was found to have sensitivity greater than 96% relative to virus culture for the diagnosis of H1N1 influenza. High accuracy of RT-PCR has also been reported for MERS (21). On the other hand, low accuracy has been reported for detection of SARS with real-time RT-PCR (22,23), although rates of detection were improved with the refinement of laboratory methods (24).

In the current COVID-19 pandemic, it has been a great boon to have had the rapid development of several versions of RT-PCR diagnostic tests that target the detection of different genes from the viral RNA. Laboratory testing has shown that at least one version of the RT-PCR assay can detect viral loads as small as 3.2 RNA copies per reaction (4) and that it does not cross-react to other known coronaviruses, particularly when the primer for the assay is well-chosen (7). However, there is widespread doubt about how well the tests work in practice (12,13). One source of error arises from the uncertain distribution of the virus in the body at various times during the COVID-19 disease trajectory (11). Comparisons of specimens from nasal and throat

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

swabs indicate better sensitivity in nasal swabs and diminished sensitivity in throat swabs, particularly after the first few days of disease onset (10). The variation in the severity of the viral infection between subjects is another source of error; milder infections are more likely to escape detection. Other sources of error include sample collection, storage and transportation errors, such as collecting a low volume of fluid in swabs and depletion of the sample. Laboratory errors during assay processing are possible too.

The sensitivity and specificity of diagnostic testing using RT-PCR for COVID-19 that were estimated in this study may be considered to provide the cumulative impact of the various possible sources of error. It is clear that the sensitivity of the test is its weakest aspect while the specificity appears to be very good. Between 15% and 29% of COVID-19 cases may have gone undetected by the RT-PCR diagnostic test designed by China CDC that was implemented with TaqMan One-Step RT-PCR kits. As far as the medical practitioner is concerned, the predictive values of the diagnostic test are of utmost importance. For COVID-19, if a medical practitioner suspected that there was a 50% chance that a patient had the disease, a subsequent positive RT-PCR test would increase that chance to 98.4%. On the other hand, after a negative RT-PCR test the patient still has 18.4% chance of the disease. A second confirmatory negative test would be needed to bring the chance of disease below 5%.

It is possible that some of the diagnostic errors were mitigated by the actions of medical professionals who might have taken a critical view of negative test results for symptomatic patients. Nevertheless, the false negative rate is still likely to be among the main reasons for the difficulty in controlling the breakout in its early stages. The problem of false negatives implies that public health measures that rely on singling out and isolating the cases of COVID-19 are unlikely to be successful on large scales. For instance, if the true prevalence is 1% in a population, testing would miss approximately 22 cases of COVID-19 for every 10,000 people tested. A highly transmissible virus can continue to propagate through the misdiagnosed cases.

The proportion of throat swab specimens that were positive for SARS-CoV-2 in RT-PCR tests conducted on patients with confirmed COVID-19 have been reported as being a mere 32% (11) and almost twice as much – 60% in severe cases and 61% in mild cases – in another study (10). Neither of those values is equivalent to sensitivity as defined in this study because multiple specimens were drawn from a smaller set of participants in the mentioned studies. Nonetheless, it seems reasonable to conclude that the sensitivity estimated in this study is higher than what was suggested by the mentioned studies. A possible reason might be that viral loads may have been higher for data collected in the epicenter of the pandemic. On the other hand, it is worth noting that the sensitivity of RT-PCR for SARS-CoV-2 that was estimated in this study is in close alignment with the value of 0.80 reported for the sensitivity of detection of SARS with RT-PCR (24). Perhaps the similarity is not too surprising since the genomes of the two viruses have been reported to be 82% similar (25).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

The primary limitation of this study is that the estimated sensitivity and specificity apply to the particular version of the RT-PCR test that was urgently created (3) and that was being used in Wuhan, China, during January and February, 2020. Laboratories around the world reacted rapidly to the pandemic and created their own versions of the RT-PCR test, as well as tests of other types. It may be expected that experimentation was done with protocols and procedures that resulted in changes in the performance of RT-PCR tests that were developed later. Another limitation of the study is that it is a retrospective study based on probabilistic knowledge of diagnostic errors. A study that is designed to compare the diagnostic accuracy of RT-PCR with a better gold standard method would be able to provide more definitive estimates and narrower confidence intervals. Data about the severity of infections of sampled patients and measures of viral load that were found in the RT-PCR tests, such as cycle threshold, were not available, which is another limiting factor of this study.

Conclusions

The diagnostic sensitivity and specificity of the RT-PCR test for COVID-19 were reconstructed from data on 1014 patients in Wuhan, China. Uncertainty that arose from incomplete knowledge of the joint distributions of test results and disease status was quantified with a modified Bayesian analysis, along with the quantification of uncertainty due to sampling error. The results indicated that the RT-PCR test administered via throat swabs had a conspicuous rate of false negative results, likely missing between 15% and 29% of patients with COVID-19. For any patient who is suspected to have COVID-19 with higher than a roughly 1-in-5 chance, at least two confirmatory negative RT-PCR tests would be necessary to reduce the likelihood of disease below 5%. The limitation of the study findings is that they apply to one version of the RT-PCR diagnostic test for COVID-19 that was developed and distributed urgently by China CDC. Study findings may not generalize to other versions of the RT-PCR test that are being used in diverse geographic regions.

Acknowledgements

I wish to thank Stanley Cron for his review and feedback on the first draft of this manuscript.

References

1. WHO Coronavirus [Internet]. [cited 2020 Apr 16]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019
2. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579(7798):265–9.
3. Chinese National Institute for Viral Disease Control and Prevention [Internet]. Primers and probes for detecting the novel coronavirus with RT-PCR. [cited 2020 Apr 23]. Available from: http://ivdc.chinacdc.cn/kyjz/202001/t20200121_211337.html

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

4. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance. 2020 Jan 23;25(3):2000045.
5. Reusken CBEM, Broberg EK, Haagmans B, Meijer A, Corman VM, Papa A, et al. Laboratory readiness and response for novel coronavirus (2019-nCoV) in expert laboratories in 30 EU/EEA countries, January 2020. Eurosurveillance. 2020 Feb 13;25(6):2000082.
6. Pang J, Wang MX, Ang IYH, Tan SHX, Lewis RF, Chen JI-P, et al. Potential Rapid Diagnostics, Vaccine and Therapeutics for 2019 Novel Coronavirus (2019-nCoV): A Systematic Review. J Clin Med. 2020 Feb 26;9(3).
7. Chan JF-W, Yip CC-Y, To KK-W, Tang TH-C, Wong SC-Y, Leung K-H, et al. Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-polymerase chain reaction assay validated in vitro and with clinical specimens. J Clin Microbiol [Internet]. 2020 Mar 4 [cited 2020 Apr 7]; Available from: https://jcm.asm.org/content/early/2020/03/27/JCM.00310-20
8. WHO COVID-19 Technical Guidance [Internet]. [cited 2020 Apr 23]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance
9. Pfefferle S, Reucher S, Nörz D, Lütgehetmann M. Evaluation of a quantitative RT-PCR assay for the detection of the emerging coronavirus SARS-CoV-2 using a high throughput system. Eurosurveillance. 2020 Mar 5;25(9):2000152.
10. Yang Y, Yang M, Shen C, Wang F, Yuan J, Li J, et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. medRxiv. 2020 Feb 17;2020.02.11.20021493.
11. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. JAMA [Internet]. 2020 Mar 11 [cited 2020 Apr 9]; Available from: https://jamanetwork.com/journals/jama/fullarticle/2762997
12. Lazar K, Ryan A. How accurate are coronavirus tests? Doctors raise concern about 'false-negative' results. BostonGlobe.com [Internet]. 2020 Apr 2 [cited 2020 Apr 8]; Available from: https://www.bostonglobe.com/2020/04/02/nation/how-accurate-are-coronavirus-tests-doctors-raise-concern-about-false-negative-results/
13. Krumholz HM. If You Have Coronavirus Symptoms, Assume You Have the Illness, Even if You Test Negative. The New York Times [Internet]. 2020 Apr 1 [cited 2020 Apr 8]; Available from: https://www.nytimes.com/2020/04/01/well/live/coronavirus-symptoms-tests-false-negative.html
14. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology. 2020 Feb 26;200642.
15. Zadeh LA. The concept of a linguistic variable and its application to approximate reasoning—I. Inf Sci. 1975 Jan 1;8(3):199–249.
16. Gelman A, Carlin JB, Stern H, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis. 2014.
17. Johnson NL, Kotz S, Balakrishnan N. Ch. 21 Beta Distributions. In: Continuous univariate distributions. 2nd ed. New York: Wiley; 1994. (Wiley series in probability and mathematical statistics).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.24.20078949; this version posted April 29, 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 4.0 International license .

18. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: http://www.R-project.org
19. RStudio Team. RStudio: Integrated Development for R. [Internet]. Boston, MA: RStudio, Inc.; 2019. Available from: http://www.rstudio.com/
20. López Roa P, Catalán P, Giannella M, García de Viedma D, Sandonis V, Bouza E. Comparison of real-time RT-PCR, shell vial culture, and conventional cell culture for the detection of the pandemic influenza A (H1N1) in hospitalized patients. Diagn Microbiol Infect Dis. 2011 Apr;69(4):428–31.
21. Huh HJ, Kim JY, Kwon HJ, Yun SA, Lee MK, Ki CS, et al. Performance Evaluation of the PowerChek MERS (upE & ORF1a) Real-Time PCR Kit for the Detection of Middle East Respiratory Syndrome Coronavirus RNA. Ann Lab Med. 2017 Nov;37(6):494–8.
22. Peiris J, Chu C, Cheng V, Chan K, Hung I, Poon L, et al. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. The Lancet. 2003 May 24;361(9371):1767–72.
23. Lau LT, Fung Y-WW, Wong FP-F, Lin SS-W, Wang CR, Li HL, et al. A real-time PCR for SARS-coronavirus incorporating target gene pre-amplification. Biochem Biophys Res Commun. 2003 Dec 26;312(4):1290–6.
24. Poon LLM, Chan KH, Wong OK, Yam WC, Yuen KY, Guan Y, et al. Early diagnosis of SARS Coronavirus infection by real time RT-PCR. J Clin Virol. 2003 Dec 1;28(3):233–8.
25. Chan JF-W, Kok K-H, Zhu Z, Chu H, To KK-W, Yuan S, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microbes Infect. 2020;9(1):221–36.