

The epistemology of a positive SARS-CoV-2 test

Rainer J. Klement¹, Prasanta S. Bandyopadhyay²

1 Department of Radiotherapy and Radiation Oncology, Leopoldina Hospital Schweinfurt, Germany

2 Department of History & Philosophy, Montana State University, Bozeman, MT, USA

* Correspondence to:

Dr. Rainer Johannes Klement

Department of Radiotherapy and Radiation Oncology

Leopoldina Hospital Schweinfurt

Robert-Koch-Straße 10

97422 Schweinfurt

Email: rainer_klement@gmx.de

Phone: +49 9721 720 2761

Homepage: www.rainerklement.com

ORCID ID: <https://orcid.org/0000-0003-1401-4270>

Abstract

We investigate the epistemological consequences of a positive polymerase chain reaction SARS-CoV test for two relevant hypotheses: (i) V is the hypothesis that an individual has been infected with SARS-CoV-2; (ii) C is the hypothesis that SARS-CoV-2 is the cause of flu-like symptoms in a given patient. We ask two fundamental epistemological questions regarding each hypothesis: First, how much confirmation does a positive test lend to each hypothesis? Second, how much evidence does a positive test provide for each hypothesis against its negation? We respond to each question within a formal Bayesian framework. We construe degree of confirmation as the difference between the posterior probability of the hypothesis and its prior, and the strength of evidence for a hypothesis against its alternative in terms of their likelihood ratio. We find that test specificity – and coinfection probabilities when making inferences about C – were key determinants of confirmation and evidence. Tests with <87% specificity could not provide strong evidence (likelihood ratio >8) for V against $\neg V$ regardless of sensitivity. Accordingly, low specificity tests could not provide strong

evidence in favor of C in all plausible scenarios modeled. We also show how a positive influenza A test disconfirms C and provides weak evidence against C in dependence on the probability that the patient is influenza A infected given that her symptoms are not caused by SARS-CoV-2. Our analysis points out some caveats that should be considered when attributing symptoms or death of a positively tested patient to SARS-CoV-2.

Keywords

Bayesianism; Confirmation; COVID-19; Evidence; RT-qPCR

Declarations

Funding: This research received no external funding.

Conflicts of interest/Competing interests: The authors declare that they have no competing interests relevant to this research.

Availability of data and material: NA

Code availability: NA

Authors' contributions: RJK conceived of the presented idea and wrote the manuscript; PSB developed the theoretical conceptions and supervised the project. Both authors discussed the results and contributed to the final manuscript

Acknowledgments

We thank Prof. Dr. Ulrike Kämmerer for helpful discussions of the RT-qPCR method and constructive comments. We also thank Prof. Gordon Brittan Jr., Prof. Palash Sarkar and an anonymous reviewer for helpful comments on a previous version of this paper.

Introduction

Currently, the world is a state of emergency caused by the new virus called severe acute respiratory syndrome-corona virus-2, or short SARS-CoV-2. COVID-19 (coronavirus disease 2019) is the term used to describe the disease symptoms caused by the SARS-CoV-2. Symptoms of COVID-19 are typically flu-like symptoms such as fever and cough and in severe cases pneumonia, which however predominantly occurs in frail patients with other comorbidities. In general, SARS-CoV-2 appears to cause similar symptoms as other coronaviruses or influenza strains (Guan et al. 2020; Shi et al. 2020).

To prevent the virus from spreading and causing symptoms and deaths, policy makers at the helm of affairs had taken dramatic measures which included the closure of kindergartens, schools and universities, restaurants, museums and shops, prohibition of gatherings, cancellation of public events and the prohibition to leave the house without good reason. These measures were justified under the premise that SARS-CoV-2 is extremely virulent. However, the validity of this premise itself is rarely investigated. In particular, which data justify this premise? It appears that the positive testing of patients and its reporting in the media play a central role in sustaining the belief in a high virulence of SARS-CoV-2 and the notion of a pandemic. Our goal here is therefore to conduct a critical analysis of the SARS-CoV-2 testing and the conclusions that can be drawn from it based on a Bayesian account of evidence and confirmation.

Evidence, confirmation, and diagnostic testing

A comprehensive understanding of scientific hypotheses requires an understanding of scientific inference, broadly construed. However, several epistemological issues need to be distinguished in order to appreciate the proper relationship between the tenability of scientific hypotheses and inference. We will discuss the significance of these issues/questions by borrowing an insight from Richard Royall (Royall 1997; Royall 2004). Our approach provides a unified Bayesian response to three questions posed by Royall.

Consider two hypotheses: V , stating that a patient is infected with the SARS-CoV-2, and $\neg V$, its denial. Assume that a SARS-CoV-2 test comes out positive. Based on this simple scenario, one could pose at least three types of question that underline the epistemological issues at stake, following an insight from Royall (1997):

- (i) Given the positive test result, what should we *believe* about V and to what degree?
- (ii) Does the positive test result provide strong *evidence* for V against its alternative $\neg V$?
- (iii) Given the positive test result, what should we do?

We call the first question the belief or confirmation question, the second the evidence question and the third the decision question. These three questions are pre-theoretical and statistical paradigm-neutral; yet they require some statistical/probabilistic tools for their articulation. Here, we will confine ourselves to the first two questions and just briefly touch upon the decision question in the Discussion section.

We have developed two distinct accounts to answer the first two types of questions (Bandyopadhyay and Brittan 2006; Bandyopadhyay et al. 2016). The first is an account of belief/confirmation, the second of evidence. Our two accounts of belief/confirmation and evidence naturally fulfill the need to have both notions of degrees-of-belief and degrees-of-support pointed out by Hawthorne (2005), which not only avoids the old evidence problem but also resolves a number of other philosophical paradoxes (Bandyopadhyay et al. 2016).

For Bayesians, degrees of belief need to be fine-grained (Ramsey 1926). A satisfactory Bayesian account of confirmation, according to us, should be able to capture this notion of degree of belief. In formal terms:

$$D \text{ confirms } H \text{ to a greater degree if and only if } P(H|D) > P(H), \quad (1)$$

where D denotes the data and H the hypothesis. Confirmation becomes strong or weak depending on how great the difference is between the posterior probability, $P(H|D)$, and the prior probability of the hypothesis, $P(H)$. $P(H|D)$ represents an agent's degree of belief in the hypothesis after the data are accumulated. $P(H)$ stands for an agent's degree of belief in the hypothesis before the data for the hypothesis have been acquired.

While the account of confirmation is concerned with belief in a single hypothesis, our account of evidence compares the merits of two simple statistical hypotheses, H_1 and H_2 (or $\neg H_1$) relative to the data D , some auxiliaries, and background information¹. Bayesians use the Bayes factor to make this comparison, while others use the likelihood ratio (LR) or other functions designed to measure evidence. For simple statistical hypotheses/models with no free parameters², as is the case in diagnostic testing that is treated here, the Bayes factor and the LR are identical, and capture the bare essentials of an account of evidence without any appeal to prior probability. The LR in favor of H_1 over H_2 is

$$LR = P(D|H_1) / P(D|H_2) \quad (2)$$

¹ The dependence on auxiliaries and background information will not be explicitly stated in the equations.

² See Bandyopadhyay et al. (1996) for a general model selection case when models have adjustable parameters.

The data D constitute evidence for H_1 against H_2 if and only if their LR is greater than one. Note that, if $1 < LR \lesssim 8$, then D is often said to provide weak to moderate evidence for H_1 against H_2 , while when $LR \gtrsim 8$, D provides strong evidence (Kass and Raftery 1995; Bandyopadhyay et al. 2016). This is also the cut-off and characterization of evidence we use in this work.

Sensitivity and specificity

In the following, we assume that testing for SARS-CoV-2 is based on reverse transcriptase – quantitative polymerase chain reaction (RT-qPCR), a biomedical testing procedure that is routinely applied for detecting and quantifying RNA. Such tests are able to detect very few numbers of nucleic acid molecules – in this case RNA – by amplifying a target nucleic acid sequence several million fold. Because RNA cannot serve as a template for PCR, the RNA template is first reversely transcribed (RT) into complementary DNA (cDNA)³ which is then exponentially amplified in a PCR over many cycles (Bustin 2000). In quantitative (also called real time) PCR one can calculate the number of DNA molecules of the amplified sequence that were initially present in the sample (Kubista et al. 2006, 96). However, for such a calculation to be reliable several operating standards should be followed closely, else sensitivity (true positive rate) and specificity (true negative rate) may be compromised (Raymaekers et al. 2009).

Upon the spread of the SARS-CoV-2 in Wuhan, China, new RT-qPCR tests have been rapidly developed. For example, Corman et al. created a test based on the closely related SARS-CoV from 2003 and got their paper published in January 2020 only two days after submission (Corman et al. 2020). Nowadays, a multitude of more or less defined commercial and laboratory "in house" tests are used, often without mandatory guidelines and adequate validation. While some of these tests have shown 100% specificity in independent validation studies (Nalla et al. 2020), others have been found to yield a significant percentage of false-negative results. For example, an early validation study of Chinese SARS-CoV-2 RT-qPCR tests revealed a false positive rate of almost 50% or higher (Zhuang et al. 2020) – however, the study was retracted for unknown reasons soon after ahead-of-print publication. In a German inter-laboratory validation study ("Ringversuch") of many commercially available and in-house RT-qPCR tests, a total of 67 out of 983 SARS-CoV-2-negative samples containing the human coronavirus HCoV 229E were classified as "positive", yielding an average false-positive rate of 6.8% (Zeichhardt and Kammel 2020). Another German laboratory found some commercial primers and probes to be contaminated with nucleotides, resulting in up to 7.3% false positive results from SARS-CoV-2-negative human throat swabs and up to 17% false positive results from swabs taken from cattle (Wernike et al. 2020). Finally, the test used by the US Center for

³ To provide RT enzymes a starting point for synthesis, specific primers flanking the RNA sequence to be amplified are used. Their choice is crucial for estimating the RNA copy number (Bustin 2000).

Disease Control appears to have specificity problems, as 3/10 reference samples classified as positive by RT-qPCR did not contain SARS-CoV-2 (Lee 2020).

SARS-CoV-2 RT-qPCR tests may also have problems with sensitivity in clinical routine application. The reason is that sensitivity depends on both the site from which a sample is obtained and the time relative to symptom onset (Sethuraman et al. 2020). Xie et al. (2020) reported on five patients presenting with flu-like symptoms and radiological diagnosis of COVID-19 pneumonia, but negative RT-qPCR test results; in these patients it took multiple re-tests to obtain a positive SARS-CoV-2 detection. Similar findings were reported by Li et al. (2020) who conducted multiple tests on the same patients diagnosed with COVID-19 but observed highly variable results. Here, 13.5% (48/355) of symptomatic patients who had initially tested negative were SARS-CoV-2 positive at the second test, and 18 patients had a positive test after two consecutive negative ones. In particular tests based on the frequently taken nasopharyngeal swabs appear to be not very sensitive (Xiaojing Wu et al. 2020). In a Chinese study on patients with confirmed COVID-19, sensitivity was only 32% (126/398) for pharyngeal swabs and 63% (5/8) for nasal swabs compared to 93% (14/15) for bronchoalveolar lavage fluid (Wang et al. 2020). This is consistent with the percentage of positive test results from nasopharyngeal swabs of COVID-19 patients by Guo et al. (2020) (51.9%) and Liu et al. (2020).

In summary we conclude that both the sensitivity and specificity of SARS-CoV-2 tests are context-dependent⁴. We therefore investigate how test inferences depend on various sensitivities and specificities for which we adopt specificity $\geq 80\%$ and sensitivity $\geq 30\%$ as realistic ranges.

The base rate

Due to the novelty of the SARS-CoV-2 and the fact that most infections remain undetected the base rates are still very uncertain. The base rate is also population-specific and space-time-dependent. In Table 1 we have compiled several infection rate estimates from various contexts. While taking these as base rate estimates assumes perfect tests, we are not so much interested in the exact values, but in a crude realistic range for our modelling study. We see that in asymptomatic persons, most studies support a base rate between 0 and 5%, while high-risk populations range somewhere between 5% and 20% (counting the 37.9% obtained by Figueira et al. (2020) as an outlier). Therefore, realistic base rates are somewhere in the range 0-20% which we account for by plotting base rate on a log scale in order to better visualize test inferences for low base rates.

⁴ Context-dependence of test performances is not restricted to the new SARS-CoV-2 tests. The reason is that such tests are not experiments within closed systems, but are conducted on open systems that are highly responsive to their environment or the given context (Klement and Bandyopadhyay 2019).

Population	Time	Specimen	Infection rate	Reference
161 hospitalized children in Wuhan, China	December 1, 2019 – January 16	Nasopharyngeal swab, sputum or bronchoalveolar lavage fluid	1.2%	Jian et al. (2020)
151 close household contacts of COVID-19 patients in Taiwan	January 15 – March 18, 2020	NA	4.6%	Cheng et al. (2020)
32 suspected SARS-CoV-2 cases and 337 people repatriated from China	January – February, 2020	NA	0%	Colson et al. (2020)
9199 inhabitants of Iceland with high risk of infection	January 31 – March 31, 2020	Nasopharyngeal and oropharyngeal swabs	13.3%	Gudbjartsson et al. (2020)
1911 symptomatic health care workers from Madrid, Spain	February 24 – April 30	Nasopharyngeal swab	11.1% ^a	Suárez-García (2020)
2085 hospital healthcare workers in Madrid, Spain	March 1 – 29, 2020	Nasopharyngeal and oropharyngeal swabs	37.9% ^a	Folgueira et al. (2020)
131 patients with mild influenza-like illness from a Los Angeles medical center, USA	March 12–16, 2020	Nasopharyngeal swabs	5.3%	Spellberg et al. (2020)
783 asymptomatic repatriation passengers arriving in Greece	March 20–25, 2020	Oropharyngeal swabs	3.6–6.3%	Lytras et al. (2020)
210 asymptomatic pregnant women from New York, USA	March 22 – April 4, 2020	Nasopharyngeal swabs	13.7%	Sutton et al. (2020)
400 asymptomatic health care workers in London hospital, UK	March 23– April 26, 2020	Nasal swabs	1.1%–7.1% (decreasing every week)	Treibel et al. (2020)
919 randomly chosen individuals from Gangel, Germany	March 31– April 6, 2020	Pharyngeal swabs	3.59%	Streeck et al. (2020)
2283 randomly	March 31 –	Nasopharyngeal	0.6%	Gudbjartsson et

chosen asymptomatic inhabitants of Iceland	April 4, 2020	and oropharyngeal swabs		al. (2020)
381 hospitalized patients in Wuhan, China	April 3–15, 2020	Nasopharyngeal swabs	0.3%	Wu et al. (2020)
1021 asymptomatic resuming patients in Wuhan, China	April 3–15, 2020	Nasopharyngeal swabs	0%	Wu et al. (2020)

Table 1: SARS-CoV-2 infection rates based on positive RT-qPCR test results from various settings in chronological order. We do not provide uncertainties on these estimates stemming from binomial statistics or test imperfection, since we are only interested in a crude range that the infection rates occupy. ^aSome persons were tested more than once.

The two questions revisited

Basic inferences from a positive SARS-CoV-2 test

Let us now assume that a person⁵ has been tested positive for SARS-CoV-2. We are interested in the following hypotheses:

V: The person is infected with the SARS-CoV-2

¬V: The person is not infected with the SARS-CoV-2

Given the positive test result, what should we believe about the infection status of the person and to which degree? This question can be answered by calculating the posterior probability of V and comparing it to its prior probability.

Given T, the positive test result, we can derive the posterior probability of V from Bayes' theorem:

$$P(V|T) = [P(T|V) \times P(V)]/P(T) \quad (3)$$

$$P(T) = P(T|V) \times P(V) + P(T|\neg V) \times P(\neg V) \quad (4)$$

The left panel of Figure 1 shows the degree of confirmation of the hypothesis that the person is infected as a function of the base rate plotted on a log scale and five different assumptions for test sensitivity and specificity. Even for tests with 90% sensitivity and 95% specificity, the base rate would

⁵ It does not necessarily have to be a patient with symptoms.

have to exceed 5.6% in order to raise our degrees of belief in V to more than 50%, while for tests with only 30% sensitivity and 80% specificity, a base rate >40% would be required.

The right panel of Figure 1 shows the LR as a function of the test sensitivity for two assumed test specificities ($=1-P(T|¬V)$). For 80% specificity, one would not be justified to speak of a positive test result as providing strong evidence in favor of V, regardless of test sensitivity; more generally, this holds for any specificity falling below 87%. If test specificity is 95%, however, evidence for V would become strong for sensitivity >40%.

COVID-19 and a positive SARS-CoV-2 test

Imagine now a patient presenting with flu-like symptoms. Borghetti et al. (2020) have pointed out that patients being tested positive for SARS-CoV-2 could have co-infection with other pathogens, and listed several other viruses and bacteria that are able to cause symptoms similar to COVID-19⁶. By now, several case reports (Borghetti et al. 2020; Cuadrado-Payán et al. 2020; Xiaojing Wu et al. 2020) and cohort studies (Jian et al. 2020; Kim et al. 2020; Lin et al. 2020) have revealed that co-infections between SARS-CoV-2 and other respiratory pathogens occur. Being aware of the fact, we therefore pose the following hypothesis:

C: SARS-CoV-2 infection has caused the flu-like symptoms of the patient.⁷

Implicit in this notion of causation is a presupposition that there exists at least one biological mechanism explaining how the virus can cause flu-like symptoms (Russo and Williamson 2007). This we take as our background knowledge that we do not explicitly state in the following equations. Given that COVID-19 has many symptoms in common with other viruses or common hospital bacteria, the negation of C could be conceived as the catch-all hypothesis for all these other possible causes:

¬C: A pathogen other than SARS-CoV-2 has caused the flu-like symptoms of the patient.

Therefore:

¬C & V: SARS-CoV-2 infection is associated with but has not caused the symptoms of the patient.⁸

⁶ Among them are adenovirus, bocavirus, other coronaviruses, influenza viruses, rhinovirus or the bacteria *Chlamydomphila pneumoniae* and *Mycoplasma pneumoniae*.

⁷ This is what some media reports implicitly assume to be true given a positive test.

⁸ In principle, we can let ¬C include the case that both SARS-CoV-2 and another pathogen cause the symptoms together, so that ¬C & V means “SARS-CoV-2 infection is associated with but has not *solely* caused the symptoms of the patient”. For simplicity, however, we continue by assuming the mono-causal scenario.

Now imagine a SARS-CoV-2 test is conducted and turns out positive. We can then ask to what degree the positive test result confirms hypothesis C. To this end, we construct the simple Bayesian network model depicted in Figure 2. The SARS-CoV-2 test assesses the truth of the hypothesis V, that the patient has the SARS-CoV-2 infection. As above, T stands for a positive test result. We conceive of V as a testable consequence of the hypothesis C; the truth of C is both a necessary and sufficient condition for the truth of V, because if C is true for a given patient it deductively follows that he/she must be infected with SARS-CoV-2. From this it follows that

$$P(V|C) \equiv p = 1 > P(V|\neg C) \equiv q \quad (5)$$

$$P(C|\neg V) = P(\neg V|C) = 0 \quad (6)$$

$$P(T|C) = P(T|V) = \text{sens} \quad (7)$$

To calculate the degree of confirmation given to C by a positive test report, we need to calculate

$$P(C|T) = P(T|C) \times P(C)/P(T) = \frac{P(T|C) P(C)}{P(T|C) P(C) + P(T|\neg C) P(\neg C)} \quad (8)$$

Defining $c \equiv P(C)$, equation (8) becomes

$$P(C|T) = \frac{c}{c + (1-c) \frac{P(T|\neg C)}{P(T|C)}} \quad (9)$$

To calculate the LR in the denominator of Equation (9) in terms of sensitivity and specificity, we re-express the likelihood $P(T|\neg C)$ using the definition of conditional probability:

$$P(T|\neg C) = \frac{P(T \& \neg C)}{P(\neg C)} \quad (10)$$

Furthermore, according to the law of total probability (Pearl et al. 2016):

$$P(T \& \neg C) = P(T \& \neg C \& V) + P(T \& \neg C \& \neg V) \quad (11)$$

We apply the product rules (Pearl et al. 2016) for the Bayesian network structure shown in Figure 2 to calculate the total probabilities in Equation (11):

$$P(T \& \neg C) = P(\neg C) P(V|\neg C) P(T|V) + P(\neg C) P(\neg V|\neg C) P(T|\neg V) \quad (12),$$

By inserting Equation (12) into (10), we obtain

$$P(T|\neg C) = P(V|\neg C) P(T|V) + P(\neg V|\neg C) P(T|\neg V) \quad (13),$$

or, using the notations given in Figure 2:

$$P(T|\neg C) = q \cdot \text{sens} + (1-q)(1-\text{spec}) \quad (14)$$

We define the $LR \equiv x$ and use Equations (7) and (14):

$$x = \frac{P(T|\neg C)}{P(T|C)} = \frac{q \cdot \text{sens} + (1-q)(1-\text{spec})}{\text{sens}} \quad (15)$$

Inserting (15) into (9) we get

$$P(C|T) = \frac{c}{c + (1-c)x} \quad (16)$$

Note that $q \equiv P(V|\neg C)$ in Equation (15) is the proportion among symptomatic cases not caused by SARS-CoV-2 which in addition to the symptom-causing pathogen carry the SARS-CoV-2. To obtain a reference point and upper limit for the magnitude of q , we therefore look at percentages of symptomatic patients with confirmed non-SARS-CoV-2 respiratory pathogen infection who additionally were tested positive for SARS-CoV-2. Such co-infection rates were found to vary from 0/3380 in French patients tested since January 1 2020 (Colson et al. 2020) to 2/239 (0.8%) in Chinese children (Jian et al. 2020), 6/186 (3.2%) in Chinese adults (Lin et al. 2020) and 24/318 (7.5%) in Californian patients (Kim et al. 2020). Based on these data, we set an upper limit of q at 10%.

In Figure 3 we plot the posterior $P(C|T)$ against the prior probability $P(C)$, adopting four different values for q : 0.5%, 2.5%, 5% and 10%. It can be seen that the confirmation we gain in the hypothesis C depends on its prior, the value for q and more on the specificity of the test than its sensitivity. As q increases, the confirmation of C provided by a positive test result becomes weaker and less dependent on the test performance, in particular specificity. However, even in the scenario where 10% of patients having symptoms not caused by SARS-CoV-2 are co-infected with SARS-CoV-2, a positive test would raise our degrees of belief in C to more than 50% if the prior probability c would exceed 29%. This may be a reasonable assumption during the height of the COVID-19 epidemic in a given country.

In Figure 4 we have plotted the evidence provided by a positive test result for the hypothesis C against $\neg C$ as a function of the test sensitivity for different fixed values of specificity and q . Note that in the depicted situations a positive test always constitutes evidence for C (likelihood ratio >1). This is consistent with the theorem that data constitute evidence for a hypothesis against its mutually exclusive and jointly exhaustive alternative if and only if the data confirm the hypothesis to a certain degree (Bandyopadhyay et al. 2016). Figure 4 shows that the evidence for C against $\neg C$ is moderate to weak if test specificity is only 80%, even for 100% sensitive tests and very low q . For $q=10\%$, this is also the case for highly specific (95%) tests. In scenarios with small probability of co-infection with SARS-CoV-2 ($q=0.5\%-5\%$) a positive test is able to provide strong evidence for C . Note however, that even for $q=0.5\%$, the evidence for C against $\neg C$ is not strong if sensitivity ranges below 40%; such low

sensitivity may be characteristic of RT-qPCR tests on pharyngeal swabs (Wang et al. 2020), so that tests using these specimen may generally be considered to allow no strong evidential inferences.

COVID-19 and a positive test for influenza A

Influenza A is one of the dominating respiratory viruses responsible for causing flu-like symptoms (Nickbakhsh et al. 2019). Coinfection of influenza A and SARS-CoV-2 has been documented (Cuadrado-Payán et al. 2020; Kim et al. 2020; Nowak et al. 2020; Xiaojing Wu et al. 2020). Imagine now that instead of a SARS-CoV-2 test, a test for influenza A has been performed in a patient with flu-like symptoms and came out positive. How does this affect an agent's degree of belief in C and the evidence for C against its alternative? Using a model with the same structure as Figure 2 and similar notation as above, let T' denote the positive test result (now for influenza A) and V' the hypothesis that an influenza A infection is present. However, now we do not have a deductive relationship between C and V' as was the case in the SARS-CoV-2 test example. Instead, V' can be conceived as a testable consequence of C in the sense that $P(V'|C) < P(V'|\neg C)$.⁹ Let us define $p' \equiv P(V'|C)$ and $q' \equiv P(V'|\neg C)$. We assume $q' > p'$ because if C is true, it follows deductively that the patient must have SARS-CoV-2, so that p' is the probability of a co-infection of SARS-CoV-2 with influenza A which appears to be $\approx 1\%$ (Kim et al. 2020). In contrast, if $\neg C$ is true, the symptoms are caused by a pathogen other than SARS-CoV-2, so q' represents the sum of (i) the probability of a single influenza A infection plus (ii) the probability of co-infection of a non-SARS-CoV-2 pathogen with influenza A plus (iii) the probability of SARS-CoV-2 infection not causing the symptoms with influenza A causing the symptoms. The prevalence of influenza A infection among symptomatic patients testing negative for SARS-CoV-2 ranges from 10/845 (1.2%) (Nowak et al. 2020) over 29/1101 (2.6%) (Kim et al. 2020) to 794/3380 (23.5%) (Colson et al. 2020). In these patients, because SARS-CoV-2 was not detected, we assume that $\neg C$ was true. We also neglect any imperfections of the influenza A tests so that these fractions give us a crude lower limit for realistic estimates of q' and show that the assumption $q' > p'$ is justified. Accordingly, we set $p'=1\%$ and vary q' from 1.2% to 25%.

The LR for the case of a positive influenza A test is

$$x' = \frac{P(T'|\neg C)}{P(T'|C)} \quad (17)$$

⁹ We thereby apply a similar definition of the testable consequence as Bovens & Hartmann (2003) in their chapter on confirmation (page 90), the difference being that the probability of the consequence given that the hypothesis is false is greater than the probability of the consequence given that the hypothesis is true in our case, while it is the other way around in Bovens & Hartmann (2003).

The numerator in the LR can be re-formulated analogous to Equation (14). A similar derivation of the denominator gives

$$P(T'|C) = \frac{P(T' \& C)}{P(C)} = \frac{P(T' \& C \& V') + P(T' \& C \& \neg V')}{P(C)} = P(V'|C) P(T'|V') + P(\neg V'|C) P(T'|\neg V') \quad (18)$$

Using $q' > p'$, we thus obtain

$$x' = \frac{q' \cdot \text{sens}' + (1-q')(1-\text{spec}')}{p' \cdot \text{sens}' + (1-p')(1-\text{spec}')} = \frac{(\text{sens}' + \text{spec}' - 1)q' + 1 - \text{spec}'}{(\text{sens}' + \text{spec}' - 1)p' + 1 - \text{spec}'} > 1 \quad (19)$$

Hereby, sens' and spec' denote the sensitivity and specificity, respectively, of the influenza A test. These are better known than for the new SARS-CoV-2 test. We adopt the values derived by López Roa et al. (2011) for the influenza A RT-qPCR test which were obtained by comparison to conventional cell culture as the gold standard. These are $\text{sens}'=95.6\%$, $\text{spec}'=82.3\%$. Figure 5 plots the disconfirmation of C constituted by a positive influenza A test for five different values of q' (left panel) as well as the evidence for $\neg C$ against C as a function of q' (right panel). Two results can be read off the graph: First, the smaller the chance of having co-infection with influenza A given that the symptoms are in fact caused by SARS-CoV-2, the stronger a positive influenza A test disconfirms the hypothesis C. Second, for all assumed values of q' , a positive test for influenza A would only constitute weak evidence for the hypothesis that the symptoms are caused by a pathogen other than SARS-CoV-2 ($\neg C$) against C.

VI. Discussion

We here provided a critical investigation of the inferences that can be drawn from a positive SARS-CoV-2 RT-qPCR test result for two main hypotheses: One is that a patient is infected with SARS-CoV-2 (hypothesis V), the second that the flu-like symptoms of a patient are caused by this virus (C) and not any other pathogen ($\neg C$). The two epistemological questions we posed are: Given a positive SARS-CoV-2 test result (i) what should we believe about each hypothesis to what degree? (ii) What is the evidence for each hypothesis compared to its negation?

We found that even for tests with 90% sensitivity and 95% specificity, the base rate would have to exceed 5.6% in order to raise our degrees of believe in V to more than 50%, while for tests with only 30% sensitivity and 80% specificity, a base rate $>40\%$ would be required. Since base rates $>40\%$ are probably unrealistic (Table 1), but sensitivity around 30% has been reported for pharyngeal swabs (Wang et al. 2020), tests based on pharyngeal swab specimen would not convince a rational agent

that the tested person is infected with SARS-CoV-2. In addition, if the test would only have 80% specificity, positive tests cannot provide strong evidence ($LR > 8$) in favor of V against $\neg V$ (Figure 1, right panel). Qualitatively, the same result was obtained regarding hypothesis C (Figure 4). Thus, sample handling and contamination that might compromise specificity become crucial.

The degree of confirmation for C was found to highly depend on the test specificity and in addition on the probability that a patient has asymptomatic SARS-CoV-2 coinfection when in fact his or her symptoms are caused by another pathogen. If this probability (that we denoted as q) is low, the degree of confirmation becomes high, and vice versa, *ceteris paribus*.

Finally, we showed that a positive influenza A test would disconfirm C to a degree depending on the probability that the patient has influenza A infection given that his or her symptoms are not caused by SARS-CoV-2. As this probability, denoted q' , increases, the degree of disconfirmation also increases, *ceteris paribus*. However, a positive influenza A test provides only weak evidence against C for the realistic ranges of q' considered. This is mainly because (i) we adopted a realistic assumption about the prevalence of influenza A infections that, despite being a very frequent pathogen causing flu-like symptoms, is less than about 25%, and (ii) because the test specificity for influenza A was only 82.3%.

Our analysis points out that one should be careful in ascribing the symptoms or death of a positively tested patient to COVID-19, if the possibility exists that the disease has been caused by another pathogen. To rule out the second possibility, one would have to test for all other possible symptom causes, which in practice is rarely attempted, and complicated by the fact that such tests are also not 100% sensitive.

Since COVID-19 is a novel disease and most tests used in practice have not been adequately validated, there are still many uncertainties associated with basic statistical quantities that we used in our modeling. We tried to account for these uncertainties by assuming several plausible values for the variables in our modeling. Yet, these uncertainties pose the major limitation of this work. In particular, the probability of having coinfection with SARS-CoV-2 when in fact the symptoms are caused by another pathogen is crucial to the inference that can be made from a positive SARS-CoV-2 test and should be investigated in future studies.

We now briefly turn to the third epistemological question that we posed in the beginning, but so far have neglected: Given a positive test result, what should we do? We can conclude that unless one is certain that the test has a high specificity, clinical decision making should not be based solely on such tests. To test specificity, Wernike et al. (2020) recommended to pre-test each batch of PCR reagents at least 50 times with negative control samples. Furthermore, given that COVID-19 might show a

seasonality similar to other human coronaviruses (Olofsson et al. 2011; Nickbakhsh et al. 2019; Monto et al. 2020) and that similar symptoms might be caused by certain bacteria (Lin et al. 2020), the possibility of other viral or bacterial infections should always be considered. That this approach is feasible has been shown by initial studies (Jian et al. 2020; Kim et al. 2020; Lin et al. 2020; Nowak et al. 2020). In general, policy makers and the media should recognize the limitations of the new SARS-CoV-2 tests and consider the possibility that deceased patients who were tested positive for this virus might only have died with but not because of it.

References

- Bandyopadhyay, Prasanta S., and Gordon G. Brittan. 2006. Acceptability, evidence, and severity. *Synthese* 148: 259–293. <https://doi.org/10.1007/s11229-004-6222-6>.
- Bandyopadhyay, Prasanta S, Gordon Brittan Jr., and Mark L Taper. 2016. *Belief, Evidence, and Uncertainty: Problems of Epistemic Inference*. 1st ed. Basel: Springer International Publishing. <https://doi.org/10.1007/978-3-319-27772-1>.
- Borghetti, Alberto, Elena Ciccullo, ArturoVisconti, Enrica Tamburrini, and Simona Di Giambenedetto. 2020. COVID-19 diagnosis does not rule out other concomitant diseases. *European Journal of Clinical Investigation*: e13241. <https://doi.org/10.1111/eci.13241>.
- Bovens, Luc, and Stephan Hartmann. 2003. *Bayesian Epistemology*. 1st ed. New York: Oxford University Press.
- Bustin, S. A. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* 25: 169–193. <https://doi.org/10.1677/jme.0.0250169>.
- Cheng, Hao-Yuan, Shu-Wan Jian, Ding-Ping Liu, Ta-Chou Ng, Wan-Ting Huang, Hsien-Ho Lin, and Taiwan COVID-19 Outbreak Investigation Team. 2020. Contact Tracing Assessment of COVID-19 Transmission Dynamics in Taiwan and Risk at Different Exposure Periods Before and After Symptom Onset. *JAMA Internal Medicine*. <https://doi.org/10.1001/jamainternmed.2020.2020>.
- Colson, Philippe, Bernard La Scola, Vera Esteves-Vieira, Laetitia Ninove, Christine Zandotti, Marie-Thérèse Jimeno, Céline Gazin, et al. 2020. Letter to the editor: Plenty of coronaviruses but no SARS-CoV-2. *Eurosurveillance* 25: 6–7. <https://doi.org/https://doi.org/10.2807/1560-7917.ES.2020.25.8.2000171>.
- Corman, Victor M., Olfert Landt, Marco Kaiser, Richard Molenkamp, Adam Meijer, Daniel Kw Chu,

- Tobias Bleicker, et al. 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* 25: 1–8. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>.
- Cuadrado-Payán, Elena, Enrique Montagud-Marrahi, Manuel Torres-Elorza, Marta Bodro, Miquel Blasco, Esteban Poch, Alex Soriano, and Gaston J. Piñeiro. 2020. SARS-CoV-2 and influenza virus co-infection. *The Lancet* 395: e84. [https://doi.org/10.1016/S0140-6736\(20\)31052-7](https://doi.org/10.1016/S0140-6736(20)31052-7).
- Folgueira, Maria Dolores, Carmen Munoz-Ruiperez, Miguel Angel Alonso-Lopez, and Rafael Delgado. 2020. SARS-CoV-2 infection in Health Care Workers in a large public hospital in Madrid, Spain, during March 2020. *medRxiv*: 2020.04.07.20055723. <https://doi.org/10.1101/2020.04.07.20055723>.
- Guan, Wei-Jie, Zheng-Yi Ni, Yu Hu, Wen-Hua Liang, Chun-Quan Ou, Jian-Xing He, Lei Liu, et al. 2020. Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England Journal of Medicine*. <https://doi.org/10.1056/NEJMoa2002032>.
- Gudbjartsson, D.F., A. Helgason, H. Jonsson, O.T. Magnusson, P. Melstad, G.L. Norddahl, J. Saemundsdottir, et al. 2020. Spread of SARS-CoV-2 in the Icelandic Population. *The New England Journal of Medicine*. <https://doi.org/10.1056/NEJMoa2006100>.
- Guo, Li, Lili Ren, Siyuan Yang, Meng Xiao, De Chang, Fan Yang, Charles S. Dela Cruz, et al. 2020. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). *Clinical Infectious Diseases*. <https://doi.org/10.1093/cid/ciaa310>.
- Hawthorne, James. 2005. Degree-of-belief and degree-of-support: Why Bayesians need both notions. *Mind* 114: 277–320. <https://doi.org/10.1093/mind/fzi277>.
- Jian, Shupeng, Panpan Liu, Ge Xiong, Zhaohui Yang, Ming Wang, Yan Li, and Xue-jie Yu. 2020. Coinfection of SARS-CoV-2 and multiple respiratory pathogens in children. *Clinical Chemistry and Laboratory Medicine*. <https://doi.org/https://doi.org/10.1515/cclm-2020-0434>.
- Kass, Robert E, and Adrian E Raftery. 1995. Bayes Factors. *Journal of the American Statistical Association* 90: 773–795.
- Kim, David, James Quinn, Benjamin Pinsky, Nigam H. Shah, and Ian Brown. 2020. Rates of Co-infection between SARS-CoV-2 and Other Respiratory Pathogens. *JAMA - Journal of the American Medical Association* 24: 4–5. <https://doi.org/10.1001/jama.2020.6266>.
- Klement, Rainer J., and Prasanta S. Bandyopadhyay. 2019. Emergence and Evidence: A Close Look at Bunge’s Philosophy of Medicine. *Philosophies* 4: 50.
- Kubista, Mikael, José Manuel Andrade, Martin Bengtsson, Amin Forootan, Jiri Jonák, Kristina Lind,

- Radek Sindelka, et al. 2006. The real-time polymerase chain reaction. *Molecular Aspects of Medicine* 27: 95–125. <https://doi.org/10.1016/j.mam.2005.12.007>.
- Lee, Sin Hang. 2020. Testing for SARS-CoV-2 in cellular components by routine nested RT-PCR followed by DNA sequencing. *International Journal of Geriatrics and Rehabilitation* 2: 69–96.
- Li, Yafang, Lin Yao, Jiawei Li, Lei Chen, Yiyan Song, Zhifang Cai, and Chunhua Yang. 2020. Stability Issues of RT-PCR Testing of SARS-CoV-2 for Hospitalized Patients Clinically Diagnosed with COVID-19. *Journal of Medical Virology*. <https://doi.org/10.1002/jmv.25786>.
- Lin, Dachuan, Lei Liu, Mingxia Zhang, Yunlong Hu, Qianting Yang, Jiubiao Guo, Yongchao Guo, et al. 2020. Co-infections of SARS-CoV-2 with multiple common respiratory pathogens in infected patients. *Science China Life Sciences* 63: 606–609. <https://doi.org/10.1007/s11427-020-1668-5>.
- Liu, Rui, Huan Han, Fang Liu, Zhihua Lv, Kailang Wu, Yingle Liu, Yong Feng, and Chengliang Zhu. 2020. Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020. *Clinica Chimica Acta* 505. Elsevier: 172–175. <https://doi.org/10.1016/j.cca.2020.03.009>.
- López Roa, Paula, Pilar Catalán, Maddalena Giannella, Darío García de Viedma, Virginia Sandonis, and Emilio Bouza. 2011. 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. *Diagnostic Microbiology and Infectious Disease* 69. Elsevier Inc.: 428–431. <https://doi.org/10.1016/j.diagmicrobio.2010.11.007>.
- Lytras, Theodore, George Dellis, Anastasia Flountzi, Sophia Hatzianastasiou, Georgia Nikolopoulou, Katerina Tsekou, Zafiris Diamantis, et al. 2020. High prevalence of SARS-CoV-2 infection in repatriation flights to Greece from three European countries. *Journal of Travel Medicine*. <https://doi.org/10.1093/jtm/taaa054>.
- Monto, Arnold S, Peter M DeJonge, Amy P Callear, Latifa A Bazzi, Skylar B Capriola, Ryan E Malosh, Emily T Martin, and Joshua G Petrie. 2020. Coronavirus Occurrence and Transmission Over 8 Years in the HIVE Cohort of Households in Michigan. *The Journal of Infectious Diseases*. <https://doi.org/10.1093/infdis/jiaa161>.
- Nalla, Arun K., Amanda M. Casto, Meei-Li W. Huang, Garrett A. Perchetti, Reigran Sampoleo, Lasata Shrestha, Yulun Wei, Haiying Zhu, Keith R. Jerome, and Alexander L. Greninger. 2020. Comparative Performance of SARS-CoV-2 Detection Assays using Seven Different Primer/Probe Sets and One Assay Kit. *Journal of Clinical Microbiology*. <https://doi.org/10.1128/jcm.00557-20>.

- Nickbakhsh, Sema, Colette Mair, Louise Matthews, Richard Reeve, Paul C.D. Johnson, Fiona Thorburn, Beatrix Von Wissmann, et al. 2019. Virus-virus interactions impact the population dynamics of influenza and the common cold. *Proceedings of the National Academy of Sciences of the United States of America* 116: 27142–27150. <https://doi.org/10.1073/pnas.1911083116>.
- Nowak, Michael D, Emilia Mia Sordillo, Melissa R Gitman, and Alberto E Paniz Mondolfi. 2020. Co-infection in SARS-CoV-2 infected Patients: Where Are Influenza Virus and Rhinovirus/Enterovirus? *Journal of Medical Virology*. <https://doi.org/10.1002/jmv.25953>.
- Olofsson, Sigvard, Robin Brittain-Long, Lars Magnus Andersson, Johan Westin, and Magnus Lindh. 2011. PCR for detection of respiratory viruses: Seasonal variations of virus infections. *Expert Review of Anti-Infective Therapy* 9: 615–626. <https://doi.org/10.1586/eri.11.75>.
- Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell. 2016. *Causal Inference in Statistics - A Primer*. 1st ed. Chichester, West Sussex: Wiley.
- Ramsey, Frank P. 1926. Truth and Probability (1931). In *The Foundations of Mathematics and other Logical Essays*, ed. R. B. Braithwaite, 1st ed., 156–198. London: Kegan, Paul, Trench, Trubner & Co.
- Raymaekers, Marijke, Rita Smets, Brigitte Maes, and Reinoud Cartuyvels. 2009. Checklist for optimization and validation of real-time PCR assays. *Journal of Clinical Laboratory Analysis* 23: 145–151. <https://doi.org/10.1002/jcla.20307>.
- Royall, Richard. 1997. *Statistical Evidence: A Likelihood Paradigm*. New York: Chapman and Hall/CRC.
- Royall, Richard. 2004. The likelihood paradigm for statistical evidence. In *The Nature of Scientific Evidence*, ed. Mark L. Taper and Subhash R. Lele, 1st ed., 119-152 (with discussion). Chicago & London: The University of Chicago Press.
- Russo, Federica, and Jon Williamson. 2007. Interpreting causality in the health sciences. *International Studies in the Philosophy of Science* 21: 157–170. <https://doi.org/10.1080/02698590701498084>.
- Sethuraman, Nandini, Sundararaj Stanleyraj Jeremiah, and Akihide Ryo. 2020. Interpreting Diagnostic Tests for SARS-CoV-2. *JAMA*.
- Shi, Heshui, Xiaoyu Han, Nanchuan Jiang, Yukun Cao, Osamah Alwalid, Jin Gu, Yanqing Fan, and Chuansheng Zheng. 2020. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *The Lancet. Infectious diseases* 20. Elsevier Ltd: 425–434. [https://doi.org/10.1016/S1473-3099\(20\)30086-4](https://doi.org/10.1016/S1473-3099(20)30086-4).
- Spellberg, Brad, Meredith Haddix, Rebecca Lee, Susan Butler-Wu, Paul Holtom, Hal Yee, and Prabhu

- Gounder. 2020. Community Prevalence of SARS-CoV-2 Among Patients With Influenza-like Illnesses Presenting to a Los Angeles Medical Center in March 2020. *JAMA*.
<https://doi.org/10.1001/jama.2020.4958>.
- Streeck, Hendrik, Bianca Schulte, Beate M Kümmerer, Enrico Richter, Tobias Höller, Christine Fuhrmann, Eva Bartok, et al. 2020. Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. *medRxiv*.
<https://doi.org/https://doi.org/10.1101/2020.05.04.20090076>.
- Suárez-García, Inés, María José Martínez de Aramayona López, Alberto Sáez Vicente, and Paloma Lobo Abascal. 2020. SARS-CoV-2 infection among healthcare workers in a hospital in Madrid, Spain. *Journal of Hospital Infection*. The Healthcare Infection Society.
<https://doi.org/10.1016/j.jhin.2020.07.020>.
- Sutton, Desmond, Karin Fuchs, Mary D'Alton, and Dena Goffman. 2020. Universal Screening for SARS-CoV-2 in Women Admitted for Delivery. *New England Journal of Medicine*: 1–2.
<https://doi.org/10.1056/nejmc2009316>.
- Treibel, Thomas A, Charlotte Manisty, Maudrian Burton, Áine McKnight, Jonathan Lambourne, João B Augusto, Xosé Couto-Parada, Teresa Cutino-Moguel, Mahdad Noursadeghi, and James C Moon. 2020. COVID-19: PCR screening of asymptomatic health-care workers at London hospital. *The Lancet* 0. Elsevier. [https://doi.org/10.1016/S0140-6736\(20\)31100-4](https://doi.org/10.1016/S0140-6736(20)31100-4).
- Wang, Wenling, Yanli Xu, Ruqin Gao, Roujian Lu, Kai Han, Guizhen Wu, and Wenjie Tan. 2020. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA - Journal of the American Medical Association* 323: 1843–1844. <https://doi.org/10.1001/jama.2020.3786>.
- Wernike, Kerstin, Markus Keller, Franz J Conraths, Thomas C Mettenleiter, Martin H Groschup, and Martin Beer. 2020. Pitfalls in SARS-CoV-2 PCR diagnostics. *Transboundary and Emerging Diseases*. <https://doi.org/10.1111/tbed.13684>.
- Wu, Xiaodong, Bo Fu, Lang Chen, and Yong Feng. 2020. Serological tests facilitate identification of asymptomatic SARS-CoV-2 infection in Wuhan, China. *Journal of Medical Virology*.
<https://doi.org/10.1002/jmv.25904>.
- Wu, Xiaojing, Ying Cai, Xu Huang, Xin Yu, Li Zhao, Fan Wang, Quanguo Li, et al. 2020. Co-infection with SARS-CoV-2 and Influenza A Virus in Patient with Pneumonia, China. *Emerging Infectious Diseases* 26: 1324–1326.
- Xie, Xingzhi, Zheng Zhong, Wei Zhao, Chao Zheng, Fei Wang, and Jun Liu. 2020. Chest CT for Typical

2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology*.

<https://doi.org/10.1148/radiol.2020200343>.

Zeichhardt, Heinz, and Martin Kammel. 2020. *Kommentar zum Extra Ringveruch Gruppe 340*

Virusgenom-Nachweis - SARS-CoV-2. Düsseldorf/Berlin.

Zhuang, GH, MW Shen, LX Zeng, BB Mi, FY Chen, WJ Liu, LL Pei, X Qi, and C Li. 2020. WITHDRAWN:

Potential false-positive rate among the “asymptomatic infected individuals” in close contacts of COVID-19 patients. *Zhonghua Liu Xing Bing Xue Za Zhi* 41: 485–488.

<https://doi.org/10.3760/cma.j.cn112338-20200221-00144>.

Figures

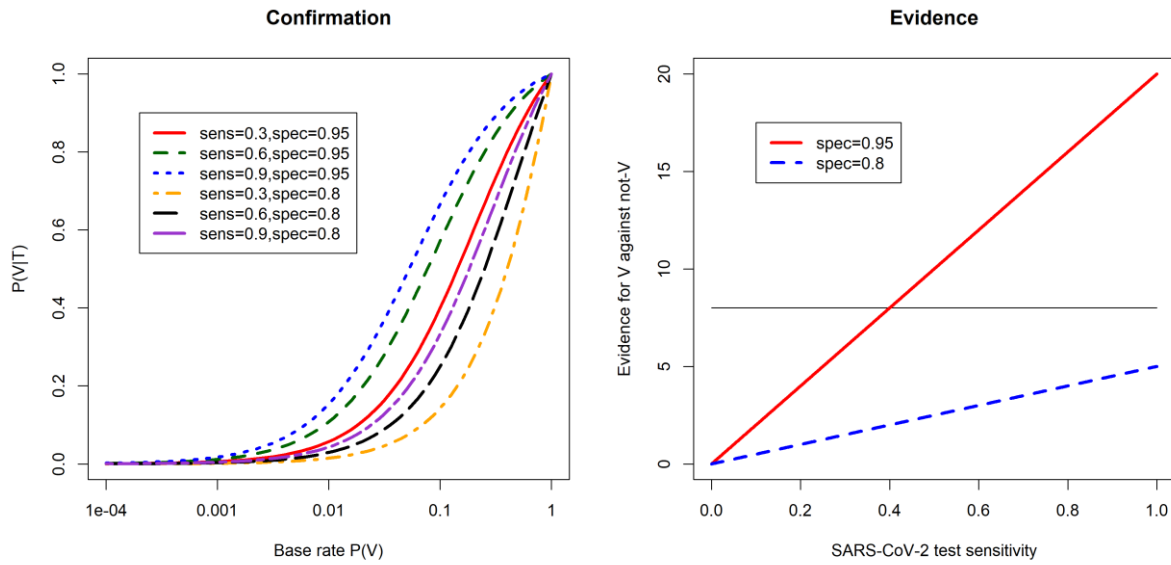


Fig. 1: Left: Posterior probability for the hypothesis V (“The tested patient is infected with the SARS-CoV-2”) as a function of the prior probability or base rate $P(V)$ plotted on a log scale. Right: Evidence measured by the likelihood ratio as a function of the test sensitivity with specificity fixed at 0.95 or 0.9, respectively. The black straight line denotes the threshold of a likelihood of 8 above which we speak of strong evidence for the hypothesis V.

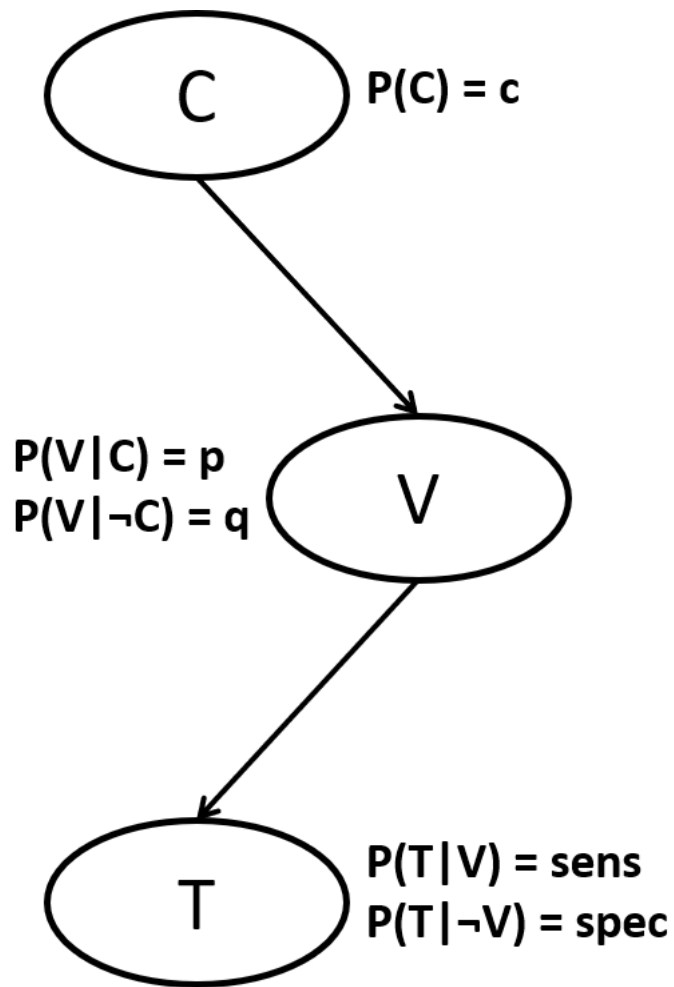


Fig. 2: Bayesian network model of the relationship between a positive test **T** and hypothesis **C**.

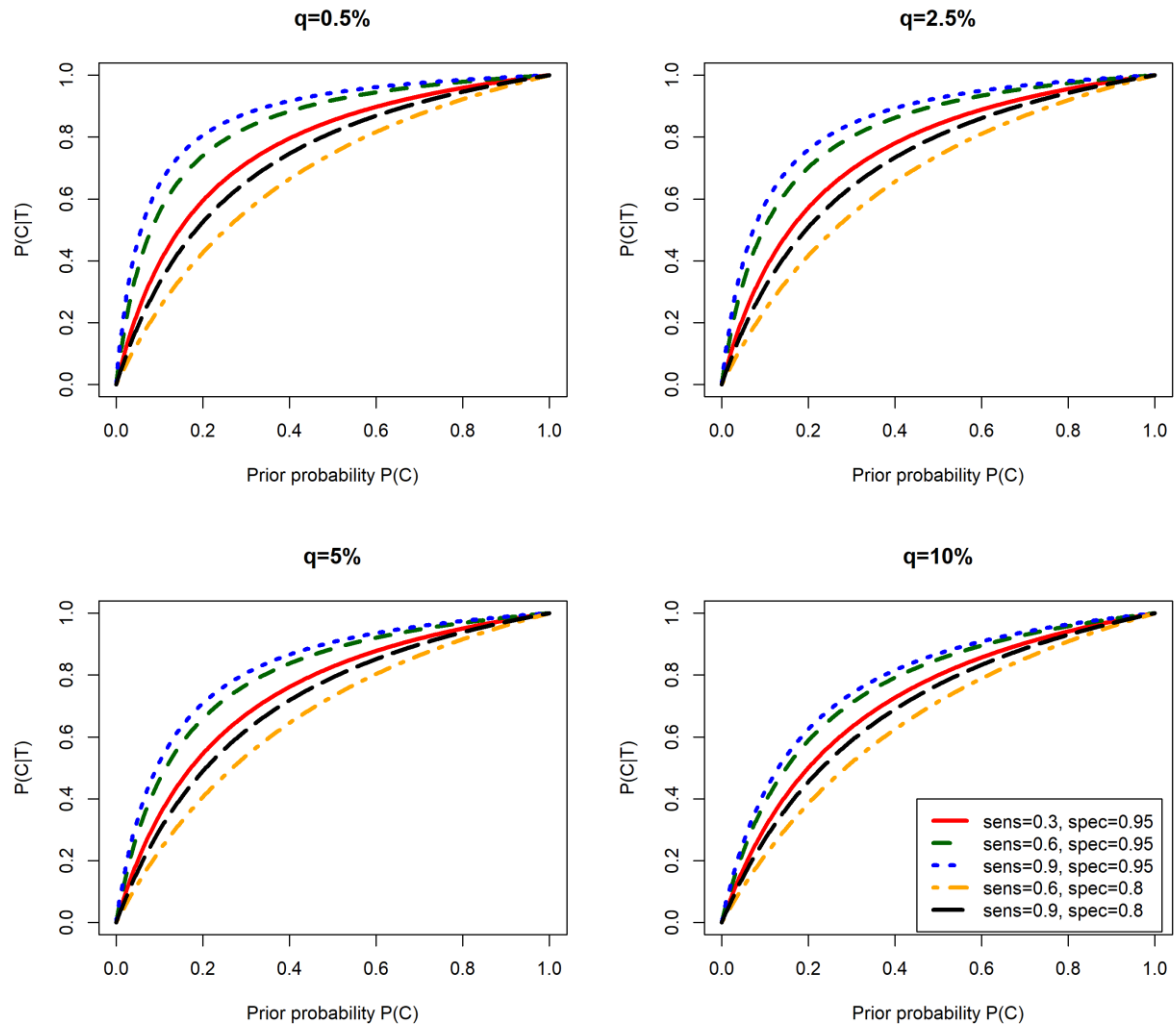


Fig. 3: Confirmation of the hypothesis C (that SARS-CoV-2 caused the COVID-19-like symptoms of a patient) by a positive test as a function of the prior probability for C, $q \equiv P(V| \neg C)$ and different test performances.

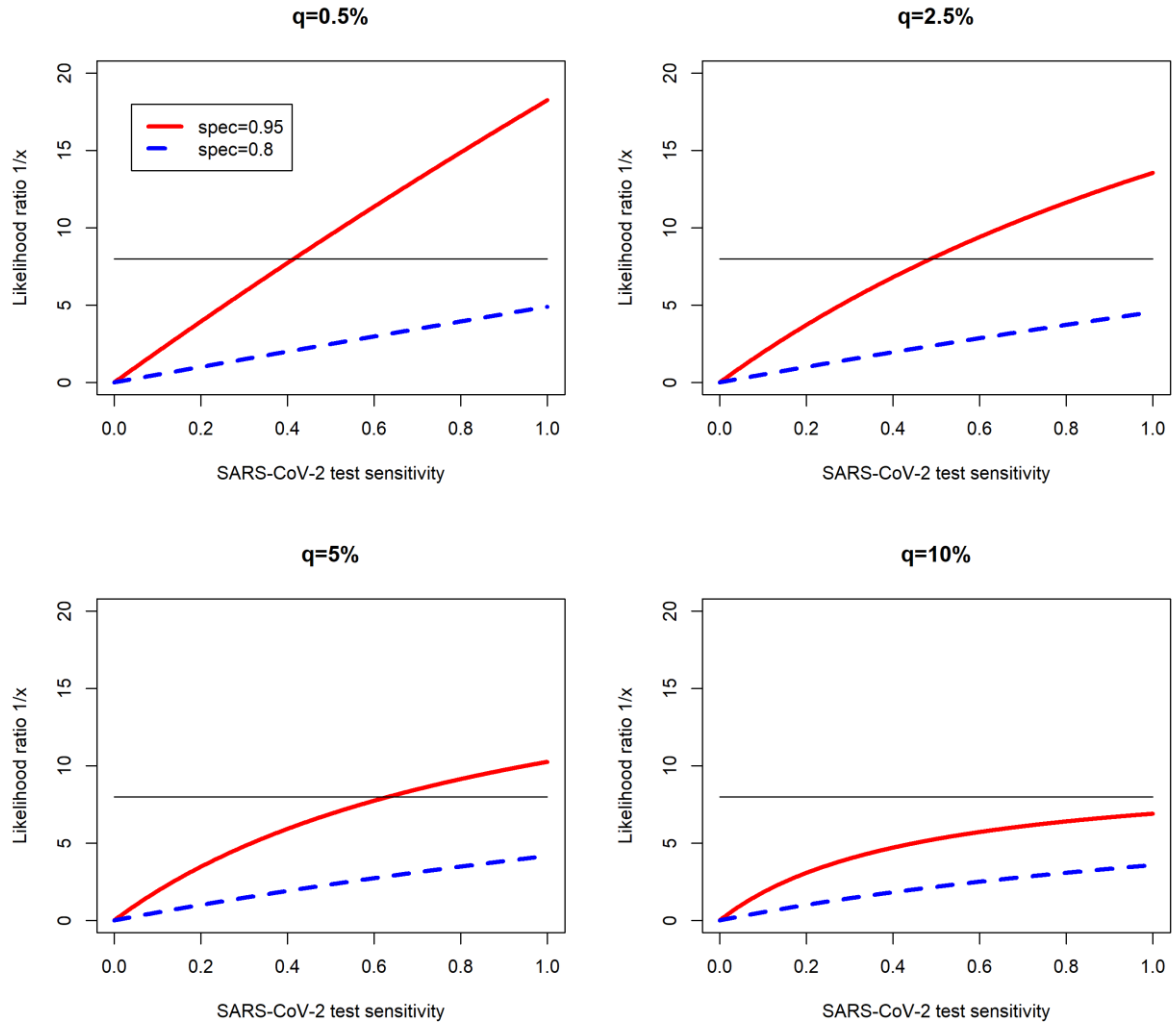


Fig. 4: Evidence for the hypothesis C against $\neg C$ given by $1/x$ (Equation 15). The evidence is plotted as a function of the test sensitivity for fixed specificity of 80% and 95%, respectively, and different values of q , the probability that a patient who has symptoms caused by a pathogen other than SARS-CoV-2 additionally has SARS-CoV-2 coinfection. The black solid line denotes the threshold of strong evidence ($1/x=8$).

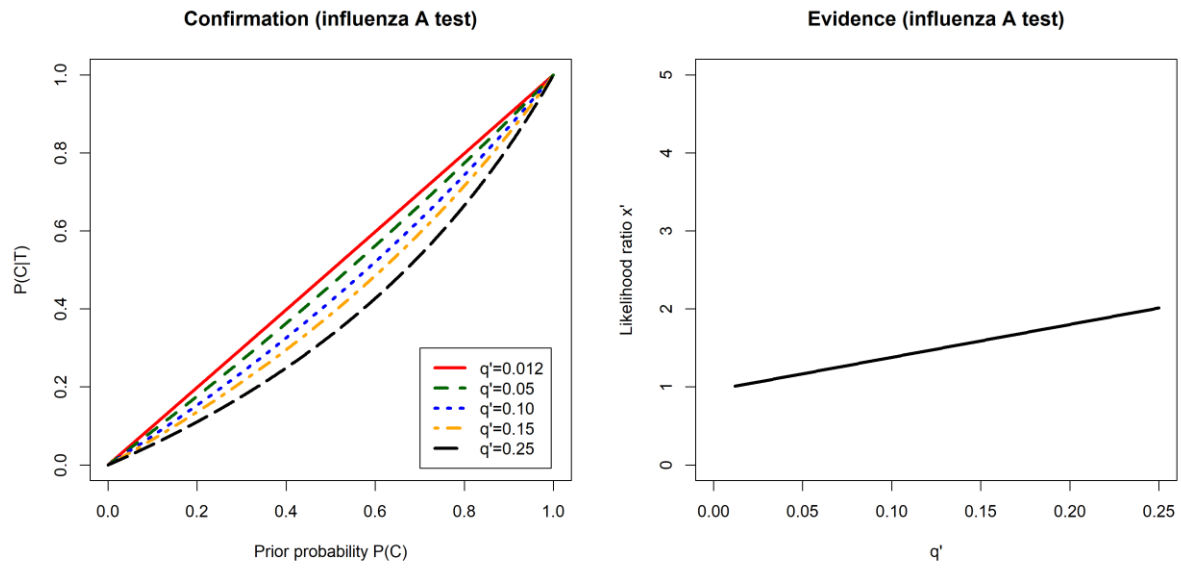


Fig. 5: (Dis-)confirmation of C and evidence of $\neg C$ against C (Equation 19) constituted by a positive influenza A test in dependence of p' , the probability of having an influenza A infection when in fact the symptoms are caused by SARS-CoV-2.