Interpreting Covid-19 Tests and the Uncertainty by Bayesian Methodology

Tomáš Karel[1] | *Prague University of Economics and Business, Prague, Czechia*

Received 9.7.2024, **Accepted** (reviewed) 10.8.2024, **Published** 13.6.2025

Abstract

The global impact of the Covid-19 pandemic has highlighted the urgent necessity for the development of rapid, effective diagnostic methods. Rapid antigen tests (RAT) have emerged as a key tool in this regard due to their speed and cost-efficiency. Nevertheless, the accurate interpretation of RAT results is challenging due to various factors, including the viral load, the quality of the sample, and the patient's status. This study demonstrates the advantages of Bayesian methods, which are capable of propagating posterior uncertainty in the form of the entire posterior distribution. It also highlights the benefits of using informative priors, which significantly reduce uncertainty in diagnostic parameters, lower false negative rates, and improve clinical decision-making. The results emphasize the need for precise interpretation of RAT results including uncertainty. Employing Bayesian simulations for posterior predictive values can reduce diagnostic errors and improve public health outcomes by upgrading the performance of RATs and explicitly propagating posterior uncertainty in clinical diagnosis, as described in this study.

Keywords	***DOI***	***JEL code***
Bayesian statistics, posterior simulations, informative prior, Covid-19	*https://doi.org/10.54694/stat.2024.36*	*C11, C13, C15*

INTRODUCTION

The Covid-19 pandemic has highlighted the urgent global need for rapid, effective and accurate diagnostic methods. Rapid antigen testing has emerged as a crucial tool in combating the spread of SARS-CoV-2 due to its time efficiency, relative simplicity, and affordability. However, accurate and critical interpretation of the results of these diagnostic tests is essential. There are many factors that influence the performance of these tests, such as viral load, sample quality, collection method and performance, sample processing, stage or timing of the test in an ongoing infection, as well as individual patient status. Tests can be subject to errors and a positive or negative result may not always indicate the true status of the infection.

An accurate interpretation of diagnostic test results necessitates the consideration of both the test characteristics and contextual factors, such as disease prevalence. The prevalence of disease can vary

[1] Department of Statistics and Probability, Faculty of Informatics and Statistics, Prague University of Economics and Business, W. Curchill Sq. 1938/4, 130 67 Prague 3, Czechia. E-mail: tomas.karel@vse.cz.

based on the indications for testing, the clinical symptoms exhibited, or the preventive screening being conducted. Diagnostic test accuracy is defined by two key parameters: sensitivity (the probability of a positive result in truly diseased individuals) and specificity (the probability of a negative result in truly non-diseased individuals). The quantification of these parameters and the associated error rates (false negatives (FN) and false positives (FP)) is of critical importance. The number of false results is contingent upon the prevalence of disease in the tested population. Failure to account for this can have significant epidemiological and economic consequences. False negative test results promote the spread of the disease in the population and increase the risk of infection. Its importance increases especially in so-called "sensitive" patient groups, gerontological, immunosuppressed or oncological patients, for whom infection status becomes a potentially life-threatening complication. On the other hand, false positive results carry the risk of erroneous placement of patients in isolation wards where the risk of infection is high and this error rate is associated with direct risk to patients.

It is crucial to consider not only the point estimates of the positive predictive value (PPV) and negative predictive value (NPV) but also the uncertainty surrounding these estimates. A significant challenge is accurately interpreting test results and reducing diagnostic uncertainty. Retesting costs are often negligible in comparison to the potential economic, health, and epidemiological consequences of misdiagnosis. Clinicians stand to benefit from an understanding of both the point estimate of PPV or NPV and the uncertainty associated with these values.

Bayesian methods provide a powerful approach to expressing PPV as a probability distribution. By employing posterior MCMC simulators, Bayesian methods can simulate marginal distributions of test parameters, such as sensitivity, specificity, and prevalence. This approach quantifies uncertainty in the form of posterior credible intervals. This article demonstrates how Bayesian analysis, incorporating external information for prevalence specification, can reduce uncertainty, achieve more accurate test interpretations, quantify diagnostic uncertainty, and reduce false positives and negatives. In the context of an emerging epidemic, our understanding of the development and manifestation of the disease, its treatment and the possibility of preventive measures is burdened with considerable uncertainty. Bayesian inference is a natural method for synthesizing dynamic, time-varying information, effectively representing and propagating uncertainty through posterior distributions of parameters.

1 LITERATURE SURVEY

The Bayesian approach has significantly enhanced the precision of pre- and post-test probabilities in the diagnostic accuracy of tests detecting SARS-CoV-2 virus, offering a powerful tool for addressing the uncertainties and biases inherent in diagnostic testing. Good et al. (2020) employed Bayesian methodology and point estimates to compare two different scenarios with different pre-test probability. Padhye (2020) applied the principles of Bayesian inference to estimate the pre-test and post-test probabilities of infection with RT-PCR, effectively accounting for variable test sensitivities and specificities, thus providing a more nuanced understanding of diagnostic outcomes. Bentley (2021) identified potential errors in SARS-CoV-2 testing, including false positives and negatives. He also proposed a Bayesian theory-based approach to correcting RT-PCR test results. Similarly, Canals and Canals (2022) employed a Bayesian framework to relate clinical presumptions to post-test probabilities, thereby enhancing the accuracy of diagnostic conclusions by different testing methods.

Furthermore, Korevaar et al. (2021) demonstrated the utility of Bayesian latent class models in improving diagnostic evaluations in the absence of a gold standard, which is of particular importance in the context of the novel coronavirus disease (Covid-19), where reference standards may be unreliable or unavailable. Another noteworthy contribution is from Williamson et al. (2023), who employed Bayesian methods to inform the determination of sample size for Covid-19 rapid tests. This illustration demonstrates the method's utility in optimizing test performance and ensuring reliable diagnostic accuracy. Moreover,

studies by Sisay et al. (2022) employed Bayesian latent class models to assess the diagnostic accuracy of RT-PCR assays in resource-limited settings, exemplifying the approach's adaptability and reliability in diverse epidemiological contexts. This approach was also corroborated by Morgan et al. (2021), who examined the impact of Bayesian adjustment on the probability estimates from diagnostic tests, thereby reinforcing the significance of Bayesian methods in enhancing diagnostic accuracy. Cao et al. (2021) further expanded the Bayesian framework by synthesizing multiple diagnostic outcomes of Covid-19 tests, demonstrating that Bayesian methods can integrate varied data sources to provide comprehensive diagnostic evaluations (Cao et al., 2021). Meanwhile, Zhang and Du (2020) employed Bayesian inference to estimate the true prevalence of Covid-19 in China, thereby providing insights into the efficacy of screening and diagnostic tests in epidemiological studies. Large study based on comparing performance of antigen tests was written by Kliegr et al. (2022). Finally, the work of Ghoshal and Tucker (2020) on Bayesian Deep Learning for the detection of Covid-19 highlights the potential of combining Bayesian methods with advanced machine learning algorithms to improve the interpretability and reliability of diagnostic models. Canyaz et al. (2022) discusses a new method using machine learning algorithms to automate the analysis of CT scans for diagnosing Covid-19 with high accuracy. Karel et al. (2022) described the economic impact of false positive test results and how Bayesian approach and retesting could minimize economic costs. Collectively, these studies demonstrate the versatility and robustness of Bayesian approaches in enhancing the diagnostic accuracy of Covid-19 tests by systematically integrating prior knowledge with new data, thereby improving the reliability and interpretability of diagnostic outcomes.

2 METHODOLOGY AND DATA

2.1 Confusion matrix

The fundamental tool for assessing the quality of diagnostic tests is the confusion matrix. This matrix conveniently displays the reliability characteristics of a test. The first column represents the positive condition (in this case, "Infected") and the second column represents the negative condition (in this case, "Healthy"). The rows display the test outcomes, which can be either positive or negative. The confusion matrix thus displays the test results as true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The values in the matrix can be presented either as absolute numbers, where the sum equals the total number of tests performed, or in relative terms, where each element of the matrix represents the probability of individual test outcomes given the actual condition.

Table 1 Confusion matrix

	Infected	**Healthy**
Negative test result	FN	TN
Positive test result	TP	FP

Source: Own construction

The primary objective of any clinical test is to maximize true positive (TP) and true negative (TN) results while minimizing false negative (FN) and false positive (FP) outcomes. The sensitivity and specificity characteristics of tests are provided by the manufacturer. However, as Kliegr et al. (2022, 2023) demonstrate, these parameters of clinical tests in field conditions may differ significantly from those obtained in laboratory settings and presented by producers.

Sensitivity (Se), also known as the true positive rate (TPR), is a measure of the test's ability to correctly identify individuals who are truly infected.

$$Se = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{1}$$

Specificity, also referred to as the true negative rate (TNR), is defined as the test's ability to correctly identify individuals who are healthy.

$$Sp = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}. \tag{2}$$

The number of false positive and false negative cases can be influenced to some extent by the design of the test and the setting of the threshold. In the case of excessively high sensitivity, the number of positive results increases, but so does the number of false positive results. Conversely, with a low threshold, the test captures only "strong signals," reducing the proportion of false positives but increasing the proportion of false negatives. The relationship between these variables can be expressed as $Se = 1 - \mathrm{FNR}$ and $Sp = 1 - \mathrm{FPR}$. In the design of clinical tests, researchers strive to find a balance and mitigate the effects of false negatives and false positives.

Another metric for the accuracy of clinical tests is accuracy. Accuracy of a diagnostic test refers to the proportion of true positive (TP) and true negative (TN) results among the total number of tested cases.

$$Accurancy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}}. \tag{3}$$

The Bayesian approach and Markov Chain Monte Carlo (MCMC) simulations are employed to model the posterior distribution of positive and negative predictive values (PPV and NPV), false positive and negative rates, respectively, using different prior distributions for prevalence. Bayesian posterior simulations were computed by JASP software (Kucharsky and Wagenmakers, 2023) and the JAGS sampler using MCMC simulations (Deppaoli et al., 2016).

A crucial aspect of modelling the posterior distributions of parameters is the specification of the prior distribution for disease prevalence, sensitivity, and specificity of the test. In this context, a beta distribution was employed due to its advantageous properties (Mossman and Berger, 2001). The Bayesian approach offers a natural means of modelling the entire posterior distribution of parameters using MCMC methods and the Gibbs sampler.

2.2 RT-PCR and antigen testing

A commonly used and efficient SARS-CoV-2 test is based on the reverse-transcription polymerase chain reaction (RT-PCR) technique. RT-PCR is often referred to as the "gold standard," and it has served as a benchmark for validating antigen tests and confirming cases against which other testing methods are measured. RT-PCR tests are considered the gold standard for detecting Covid-19 due to their high sensitivity and specificity (Filchakova, et al., 2022). These tests detect viral RNA and are highly effective in diagnosing SARS-CoV-2 virus, particularly during the early stages when the viral load is high. However, RT-PCR tests are resource-intensive, requiring specialized equipment and specialist personnel, which can limit their widespread use. (Sisay et al., 2022). Antigen tests are based on detecting specific proteins from the SARS-CoV-2 virus. They are generally more affordable and faster than RT-PCR testing and can be conducted at the point of care. Despite their lower sensitivity compared to RT-PCR, antigen tests are valuable for large-scale screening, providing rapid results that are crucial for timely isolation and treatment (Chadwick et al., 2022). Nevertheless, the accuracy of these tests can vary depending on the viral load and the stage of infection.

2.3 Bayes theorem

The confusion matrix alone does not provide direct guidance for the correct and effective interpretation of diagnostic test results. In situations of low disease prevalence, even seemingly good test characteristics in terms of sensitivity and specificity (e.g., 95%) can lead to low posterior positive predictive value (PPV) values (the probability that a person is truly infected if they test positive). This phenomenon is formalized by Bayes' theorem, which is based on the relationship of conditional probabilities and priors (sensitivity, specificity, and prevalence).

Applying and correctly understanding Bayes' rule can be challenging because the correct posterior result often conflicts with intuitive reasoning. The counterintuitive nature of the Bayesian approach arises from two factors: the base-rate fallacy, where prior probability (prevalence) is not considered (Bar-Hillel, 1980), and the psychological phenomenon known as confusion of the conditionals (Pollatsek et al., 1987). To illustrate this confusion, the probability that a randomly selected person is bleeding if attacked by a lion does not equal the probability that a randomly selected person is attacked by a lion if they are bleeding.

When evaluating results of diagnostic tests, it is crucial to consider disease prevalence alongside sensitivity and specificity. For example, if a diagnostic test has 90% sensitivity and 90% specificity, it will incorrectly classify 10% of infected individuals as healthy and 10% of healthy individuals as infected. Furthermore, if we test a group of patients in an intensive care unit with clinical symptoms and epidemiological contact in order to determine their discharge, we risk releasing 10% of patients who are still infected. Conversely, large-scale preventive testing in a population with low disease prevalence (e.g., one in a thousands) risks a significant number of false-positive results due to specificity.

The posterior probability (the positive predictive value – PPV) represents a crucial metric for the interpretation of diagnostic test results:

$$P\left(infected \mid positive\right) = \frac{P\left(positive \mid infected\right) \times P\left(infected\right)}{P\left(positive\right)}, \tag{4}$$

where *P (infected/positive)* is a conditional probability that the person is infected given the positive test result. In statistical terms, this is referred to as the posterior, while in medical terms, it is known as the post-test probability.

P (*Positive/infected*) is the conditional probability that the test result will be positive if the person is infected. In the Bayesian approach, this conditional probability is referred to as the likelihood, and in this context, it is expressedas the sensitivity of the diagnostic test. In Bayesian theory, the disease prevalence is represented by the prior probability *P*(*infected*). The marginal probability of a positive test result *P*(*positive*) is obtained using the law of total probability. A positive test result can be classified as either true positive (TP) or false positive (FP). In the context of medical diagnostics, Bayes' theorem can be expressed as follows:

$$P\left(positive\right) = P\left(positive \mid infected\right) \times P\left(infected\right) + P\left(positive \mid not\ infected\right) \times P\left(not\ infected\right). \tag{5}$$

The positive predictive value (PPV) is the posterior probability that a person is infected given a positive test result:

$$PPV = \frac{Prevalence \times Sensitivity}{Prevalence \times Sensitivity + \left(1 - Prevalence\right) \times \left(1 - Specificity\right)}. \tag{6}$$

The negative predictive value (NPV) is defined as the posterior probability that a person is not infected, given that the result of the diagnostic test was negative.

$$NPV = \frac{\left(1 - Prevalence\right) \times \left(1 - Specificity\right)}{Prevalence \times \left(1 - Sensitivity\right) + \left(1 - Prevalence\right) \times Specificity}. \tag{7}$$

In practical situations, the values of these parameters sensitivity and specificity are subject to uncertainty. The estimate of prevalence (prior distribution) is burdened with even greater uncertainty. Bayesian inference provides a method to express and explicitly propagate this uncertainty through the posterior distribution of PPV and NPV. In the Bayesian framework, parameters (sensitivity, specificity, and prevalence) are considered random variables with probability distributions conditioned on the available information and data (Baron, 1994; Mossman, Berger, 2001; Crawford et al., 2009). During any epidemic of a new disease, our knowledge and understanding regarding treatment, clinical symptomatology, and prevention evolve gradually and are accompanied by significant uncertainty. From this perspective, Bayesian inference represents a natural method for combining available information and knowledge, which evolve dynamically over time.

In binary diagnostic tests, it is possible to express the probability of a correct diagnosis, including the uncertainty associated with the test result. In diagnosing infections based on test results, it is crucial that this uncertainty is recognized, properly considered, and minimized. This article describes a methodology for quantifying and reducing the uncertainty associated with diagnostic test results. The evaluation of these diagnostic test results should not solely depend on the binary outcome; rather, clinicians and researchers should consider the magnitude of the uncertainty in the result and the test parameters and prevalence. In instances of considerable uncertainty, retesting may lead to more accurate results. Moreover, the costs associated with retesting are often negligible compared to the potential consequences of misdiagnosis.

2.4 Data

The data utilized in this article were procured from the ISIN (Czech Information System for Infectious Diseases) for the specified temporal range between December 2021 to February 2022 (Kliegr et al., 2023). For the purposes of this study, the *Panbio Covid-19 AG – Abbot Rapid Diagnostics test* was selected as it was possible to observe a sufficient number of cases within the chosen period. Furthermore, the results of the AG test were validated using the RT-PCR method on the same or the following day.

2.5 Priors

In Bayesian analysis, prior distributions are assigned to individual parameters (sensitivity, specificity, and prevalence) to represent external information. These parameters take values from 0 to 1, making the beta distribution suitable for their modeling (Bolstad, 2007; Albert, 2009; Wagenmakers et al., 2021).

2.5.1 Priors for sensitivity and specificity

Mossman and Berger (2001) recommend the use of the so-called Jeffreys prior for sensitivity and specificity in the analysis of medical tests, assigning prior beta parameters $\alpha = 0.5$ and $\beta = 0.5 \sim Beta(0.5, 0.5)$. The shape of the beta distribution assigns the highest prior probability to values close to 0 and 1. One advantage of using the beta distribution for unknown parameters is that it allows for Bayesian updating with each new data point. In this process, the prior parameters are updated with the number of successes (true positives) or failures (true negatives) as the data set grows. This property is known as a conjugate prior in Bayesian statistics (Diaconis and Ylvisaker, 1979; Raiffa and Schlaifer, 1961). When additional tests are evaluated (TP, TN, FP, FN), the prior distributions for sensitivity and specificity are updated with this new information, and the posterior distribution remains beta. The sensitivity is represented by a beta distribution with parameters (0.5 + TP, 0.5 + FN), while the specificity is represented by a beta distribution with parameters (0.5 + TN, 0.5 + FP). The posterior probability distributions of PPV and NPV, including uncertainty, can be represented by performing random sampling from the posterior marginal distributions. Repeating this process yields a very precise posterior distribution after several thousand simulations.

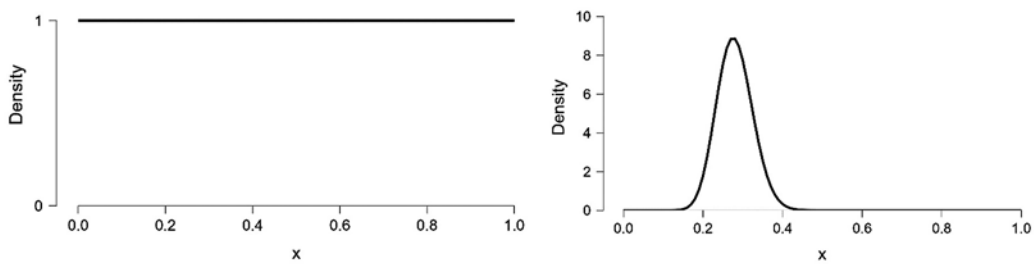
2.5.2 Priors for prevalence

This article compares the posterior distributions of PPV and NPV, including the uncertainty represented by the probability distribution of these parameters. This uncertainty, represented by the variability of the posterior distribution, is a crucial aspect in evaluating a diagnostic test. The objective of this article is to illustrate how the estimation and utilizing a prior distribution for prevalence can result in a reduction in uncertainty, as reflected by the width of the posterior 95% credible interval.

The comparison is made between a scenario with a uniform prior distribution for disease prevalence (flat prior), *Beta*(1, 1). This represents a situation where the evaluation of a clinical test result does not take disease prevalence and other external information into account. This scenario results in significant uncertainty in the posterior distribution of positive/negative predictive values.

The second scenario involves a prior distribution that takes into account the indication (motivation) for the test. In the Czech Republic, during the pandemic, indications for testing were primarily divided into preventive (screening), epidemiological, diagnostic, or clinical. In this case, the prior distribution of tested subjects was estimated based on diagnostic and clinical indications. The prior distribution for prevalence in this subgroup was estimated using data from the first 100 individuals who underwent RT-PCR testing due to clinical and diagnostic indications. This prior distribution was modeled as *Beta*(28, 72). The mean of this distribution is 0.28, which represents the expected prevalence. The standard deviation is 0.045, representing the degree of uncertainty in the Bayesian approach.

Figure 1 Noninformative and informative prior Beta distribution for prevalence



Source: Own construction

3 RESULTS

Employing an informative prior for prevalence in specific group of tested individuals significantly enhances the performance and reliability of antigen diagnostic tests compared to using a noninformative prior. This approach results in narrower posterior credible intervals across all key metrics, indicating increased precision and credibility in the test results.

Notably, the Negative Predictive Value (NPV) and accuracy show substantial improvements. The median NPV increases from 0.642 to 0.901 and accuracy rises from 0.829 to 0.902. Additionally, the widths of posterior credible intervals are reduced by 93% and 76%, respectively. This ensures more reliable identification of true negatives and enhances overall diagnostic accuracy, highlighting the effectiveness of the informative prior in clinical settings.[2]

[2] PPV – proportion of those who tested positive and are affected by the condition. P (condition = positive | test = positive). NPV – proportion of those who tested negative and are not affected by the condition. P (condition = negative | test = negative). FP – proportion of a population not affected by a condition and incorrectly tested positive. P (condition = negative ∧ test = positive). FN – proportion of a population affected by a condition and incorrectly tested negative. P (condition = positive ∧ test = negative).

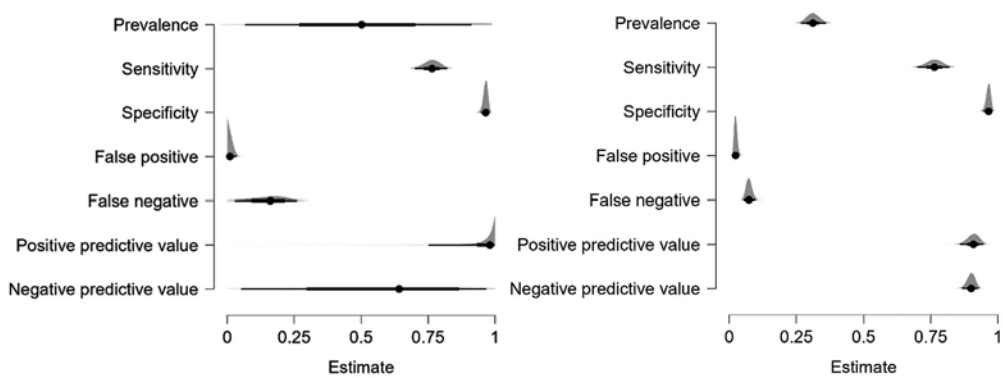
Table 2 Noninformative and informative prior comparison

	Noninformative prior			Informative prior			95% ci reduction
	Median	95% CI		Median	95% CI		
		Lower	Upper		Lower	Upper	
PPV	0.981	0.751	0.999	0.909	0.858	0.946	65%
NPV	0.642	0.051	0.967	0.901	0.867	0.927	93%
FP	0.01	0.000	0.035	0.024	0.014	0.037	34%
FN	0.161	0.029	0.26	0.073	0.054	0.097	81%
Accuracy	0.829	0.737	0.941	0.902	0.876	0.924	76%

Source: Own construction

The estimates plot presents a summary comparison of the performance metrics for a diagnostic antigen test in the context of non-informative and informative priors of prevalence. The various estimates related to the test's accuracy and reliability are displayed. This graph provides a comprehensive overview of the diagnostic performance key metrics of the antigen test in two scenarios.

Figure 2 Posterior estimates of test performance (noninformative and informative prior)



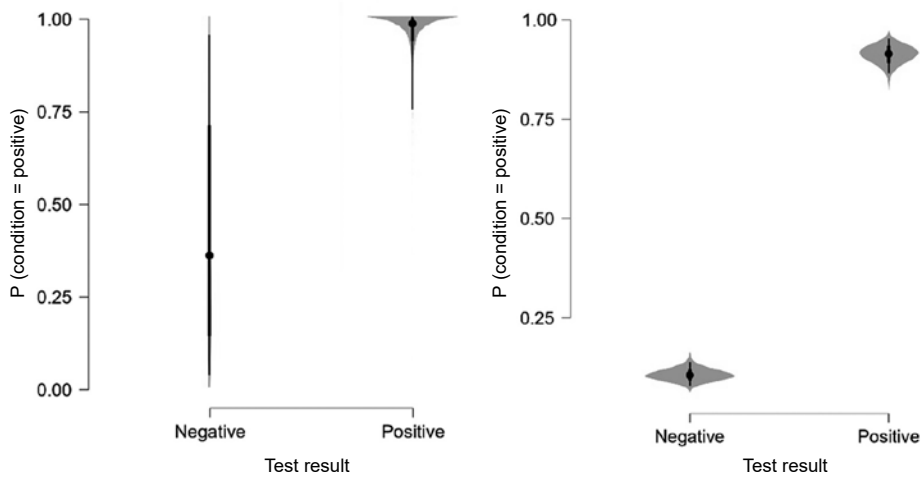
Source: Own construction, *JASP*

3.1 Posterior predictive values

Figure 3 depicts the posterior probability distribution of a condition being truly positive based on the diagnostic test result, which compares two scenarios with non-informative and informative priors for prevalence.

Using a non-informative prior for prevalence results in greater uncertainty and overlapping distributions for both negative and positive test results, reducing the reliability of the diagnostic test. In contrast, employing an informative prior significantly reduces uncertainty and provides clearer differentiation between negative and positive results, enhancing the overall effectiveness and reliability of the diagnostic test.

Figure 3 Posterior distribution (noninformative and informative priors)

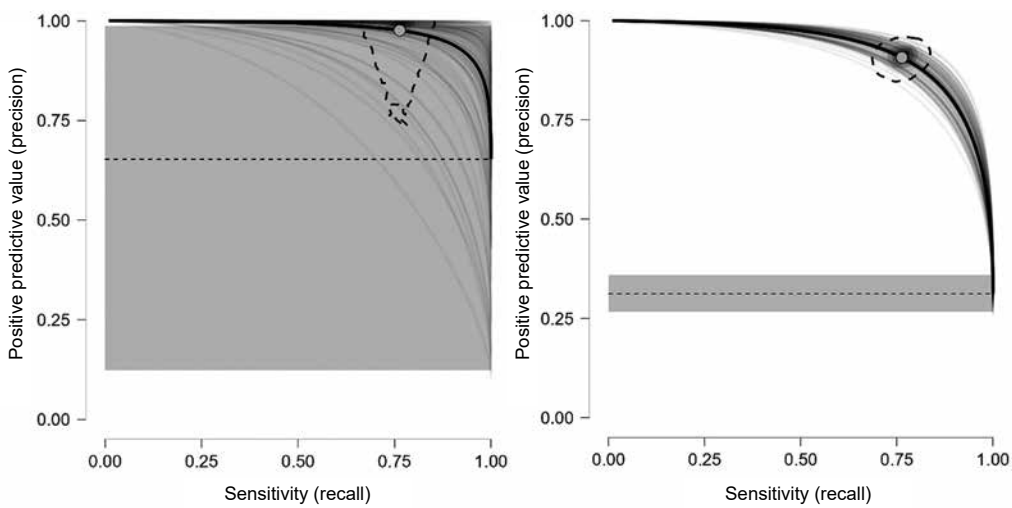


Source: Own construction, *JASP*

3.2. Precision-recall

The relationship between PPV (also known as precision) and sensitivity is another key metric for evaluating the performance of diagnostic tests. As illustrated in the graph below, the PPV uncertainty (with the dashed line representing the 95% CI for PPV) in the first situation (noninformative – flat prior for prevalence) is much larger than in the second situation with an estimated beta prior for prevalence. Using a flat prior for prevalence, the diagnostic test shows less reliable results of positive predictive values (PPV) with significant variability and uncertainty. In contrast, an informative prior results in PPV with much greater consistency and reliability. Thus, while the flat prior yields higher PPV, the informative prior enhances the test's overall stability and dependability across varying levels of sensitivity.

Figure 4 Positive predictive value – precision (noninformative and informative priors)

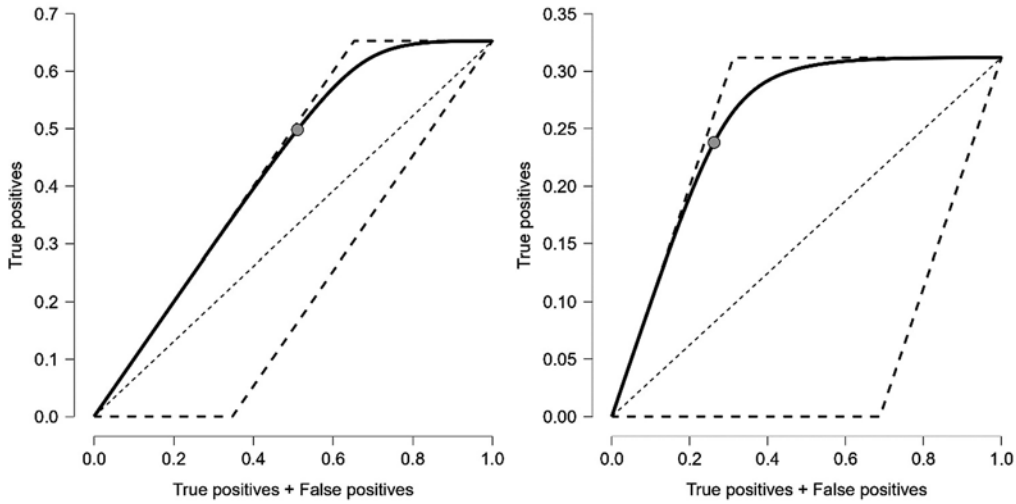


Source: Own construction, *JASP*

3.3 Total Operation Curve

A Total Operation Curve (TOC) is a further tool for evaluating the accuracy and properties of diagnostic tests, particularly for binary classification cases. In comparison to the ROC curve, it provides a more comprehensive assessment of the test's capabilities across all possible thresholds. The comparison of TOC curves for noninformative and informative priors for prevalence can be seen in the following graph.

Figure 5 TOC comparison (noninformative and informative priors)



Source: Own construction, *JASP*

The first TOC curve (non-informative prior) indicates a higher maximum true positive rate but shows lower overall test quality as it approaches the diagonal line, suggesting a higher rate of false positives. In contrast, the second TOC curve (informative prior), with a lower maximum true positive rate, demonstrates better overall performance due to its consistent growth and greater distance from the diagonal line, indicating better differentiation between true and false positives. Thus, despite the lower peak, the use of informative prior for interpretation diagnostic antigen test results is more reliable and effective overall.

CONCLUSION AND DISCUSION

In summary, the Covid-19 pandemic has highlighted the critical global need for rapid, efficient and accurate diagnostic methods. Antigen testing has emerged as a key tool in the fight against SARS-CoV-2. However, accurate interpretation of diagnostic test results remains a significant challenge. This challenge requires consideration of test characteristics such as on-field sensitivity and specificity, as well as the context, motivation behind testing, and prevalence. The accuracy of the test results is not only dependent on the type of test, with RT-PCR being considered the gold standard, but this too has its limitations. For example, in terms of the site of collection, Wang et al. (2020). found a lower positive test rate from oropharyngeal swabs (24%) compared to nasopharyngeal swabs (57%), as well as the primary interpretation of viral load. Inaccuracy, testing errors, and procedural malpractice had a significant impact on patient risk during the pandemic. The limitation of not only preventive care, but also the rapid introduction of distant diagnosis based on antigen retrieval tests (RAD) performed at home with questionable ability to perform the tests correctly, significantly increased underdiagnosis and undertreatment of patients. The lack of sensitivity and specificity of rapid antigen tests can lead to both false negatives and positives,

which significantly increases the risks of disease, its spread, and the potential risk to patients. Given the potential for Covid-19 symptoms to overlap with other diagnoses, there was a risk of delay in diagnosis of other diseases, as well as the risk of placing patients in isolation ward environments where infection rates were high. This risk became particularly important within patients from the so-called 'susceptible' population (gerontological population, immunosuppressed patients). It is therefore proposed to provide techniques to refine the interpretation of the results of these tests. In real practice, the screening use of RT-PCR is not feasible due to the procedural and time-consuming, as well as the economic aspect of these tests. Antigenic tests seem to be a suitable option, but as our statistical results have shown, it is not feasible to use only a single antigenic test without additional parameters or retesting over time, given the risks of mis-tested individuals. There are many factors that influence the performance of these tests, such as viral load, sample quality, sample processing method, stage or timing of the test in an ongoing infection, as well as individual patient status (Chaimayo et al., 2020). A study by Kliegr et al. (2022) demonstrated low on-field sensitivity across a range of commonly used RAD tests, with even the most widely used test having a sensitivity of only 48.7%. Similar results of poor sensitivity of rapid diagnostic tests were already observed during the influenza A (H1N1) pandemic. Therefore, such tests should not be recommended in isolation in clinical settings, or in widespread testing, or as a stand-alone source of information to guide further patient therapy. Our proposed model has the advantages of simplicity and the possibility of default program settings, which will also determine the level of reliability of the test results. Thus, it is possible to have a risk assessment by professional medical staff that stratifies the risk of a particular patient. Unfortunately, with every pandemic, there is a certain reduction in health care. However, it is necessary to reduce the risks from blanket measures, and refinement of testing techniques, or interpretation of test results, is an important aid.

Bayesian methods, which use posterior distributions of key parameters to more accurately represent diagnostic uncertainty, provide a robust framework for incorporating these factors. This approach improves the interpretation of positive predictive value (PPV) and negative predictive value (NPV) by accounting marginal probabilities of disease prevalence in specific subgroups, sensitivity and specificity.

The study results highlight the effectiveness of employing informative priors in Bayesian analysis. For example, the NPV median increased from 0.642 to 0.901, and accuracy improved from 0.829 to 0.902 when an informative prior was employed compared to a non-informative prior. Additionally, the posterior credible interval widths for these metrics were significantly reduced by 93% and 76%, respectively, indicating increased accuracy and decreasing uncertainty in the diagnostic test results.

This enhanced reliability is further evidenced by the reduction in false negatives (FN) from 0.161 to 0.073 and 81% reduction of posterior credible interval widths. The uncertainty in Accuracy of the diagnostic test, as demonstrated by the 95% CI, was reduced by 76%. These findings demonstrate that Bayesian approach not only aids in the accurate interpretation of test results but also contributes to reducing diagnostic errors and uncertainty, ultimately enhancing clinical decision-making and public health outcomes.

In clinical practice, the availability of rapid, critical, and more accurate diagnostic tools will not only enable early detection of infection and initiation of therapy, but also the correct indication of isolation regimens and quarantine conditions. The consequences not only of the pandemic, but also of its inevitable response, are already being observed. In March, the World Health Organization (WHO) reported a 25% increase in the prevalence of anxiety and depression. There are also efforts to examine the risk of the impact of health care restrictions during the Covid-19 pandemic. In their study, Mani and Schut (2023) observed a decrease in cardiometabolic and cancer screening. These analyses highlight the latent risks associated with health system constraints during this period. The refinement of diagnostic methods not only allows for the reduction of population stressors (fear of infection, uncertainty of false results,

and social isolation), but also allows for better management of preventive care retention and reduction of delays in elective medical interventions, contributing to a reduction in diagnostic errors and uncertainty, which ultimately improves epidemiological public health outcomes.

In evaluating diagnostic tests, researchers and clinicians should focus not only on the point estimates of PPV (post-test probabilities), but also on their entire posterior probability distribution, including the uncertainty demonstrated by the variability of this distribution. In cases where the variability of these posterior parameters is higher, it is advisable to consider retesting before making the diagnosis itself.

References

ALBERT, J. (2009). *Bayesian Computation with R* [online]. Springer. <https://doi.org/10.1007/978-0-387-92297-3>.

BAR-HILLEL, M. (1980). The base-rate fallacy in probability judgments [online]. *Acta Psychologica*, 44(3): 211–233. <https://doi.org/10.1016/0001-6918(80)90046-3>.

BARON, J. (1994). *Thinking and Deciding*. Cambridge University Press.

BENTLEY, P. M. (2021). Error rates in SARS-CoV-2 testing examined with Bayesʻ theorem [online]. *Heliyon*, 7(4): e06905. <https://doi.org/10.1016/j.heliyon.2021.e06905>.

BOLSTAD, W. M. (2007). *Introduction to Bayesian Statistics* [online]. 2nd Ed. Wiley-Interscience. <https://doi.org/10.1002/9780470181180>.

CANALS, M., CANALS, A. (2022). How accurate are radiography and computed tomography in the diagnosis of Covid-19? A Bayesian approach. *Acta Radiologica Open*.

CANAYAZ, M., ŞEHRIBANOGLU, S., ÖZDAG, R. (2022). Covid-19 diagnosis on CT images with Bayes optimization-based deep neural networks and machine learning algorithms [online]. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-022-07052-4>.

CAO, L., ZHAO, S., LI, Q., LING, L. (2021). A Bayesian method for synthesizing multiple diagnostic outcomes of Covid-19 tests [online]. *Royal Society Open Science*. <https://doi.org/10.1098/rsos.201867>.

CHADWICK, F. J., CLARK, J., CHWDHURY, S. (2022). Combining rapid antigen testing and syndromic surveillance improves community-based Covid-19 detection in a low-income country [online]. *Nature Communications*. <https://doi.org/10.1038/s41467-022-28694-6>.

CHAIMAYO, C., KAEWNAPHAN, B., TANLIENG, N., ATHIPANYASILP, N., SIRIJATUPHAT, R. et al. (2020). Rapid SARS-CoV-2 antigen detection assay in comparison with real-time RT-PCR assay for laboratory diagnosis of Covid-19 in Thailand [online]. *Virology Journal*, 17(1): 177. <https://doi.org/10.1186/s12985-020-01452-5>.

CRAWFORD, S. L., GHOSH, D., SOBEL, M. E. (2009). Bayesian Selection of Optimal Models Under Two-Stage Adaptive Designs [online]. *Journal of Statistical Planning and Inference*, 139(12): 4071–4083. <https://doi.org/10.1016/j.jspi.2009.06.012>.

DEPPAOLI, S., CLIFTON, J. P., COBB, P. R. (2016). Just Another Gibbs Sampler (JAGS): Flexible Software for MCMC Implementation [online]. *Journal of Educational and Behavioral Statistics*, 41(6): 628–649. <https://doi.org/10.3102/1076998616664876>.

DIACONIS, P., YLVISAKER, D. (1979). Conjugate Priors for Exponential Families [online]. *The Annals of Statistics*, 7(2): 269–281. <https://doi.org/10.1214/aos/117634462>.

FILCHAKOVA, O., DOSSYM, D., ILYAS AKUANYSCHEVA, T., ABDIZHAMIL, A., BUKASOV, R. (2022). Review of Covid-19 testing and diagnostic Methods [online]. *Talanta*, 244: 123409. <https://doi.org/10.1016/j.talanta.2022.123409>.

GOOD, C.B., HERNANDEZ, I., SMITH, K. (2020). Interpreting Covid-19 Test Results: a Bayesian Approach [online]. *J. Gen. Intern. Med.*, 35(8): 2490–2491. <https://doi.org/10.1007/s11606-020-05918-8>.

KAREL, T., MAZOUCH, P., FISCHER, J. (2022). The economic impact of false positivity of Covid-19 PCR testing in the Czech Republic [online]. *International Journal of Economic Sciences*, 11(1): 37-46. <https://doi.org/10.52950/ES.2022.11.1.003>.

KLIEGR, T., JARKOVSKÝ, J., JIŘINCOVÁ, H., KUCHAŘ, J., KAREL, T., TACHEZY, R. (2022). Role of population and test characteristics in antigen-based SARS-CoV-2 diagnosis, Czechia, August to November 2021 [online]. *Euro Surveill*, 27(33): 2200070. <https://doi.org/10.2807/1560-7917.ES.2022.27.33.2200070>.

KLIEGR, T., JARKOVSKÝ, J., JIŘINCOVÁ, H., KUCHAŘ, J., KAREL, T., CHUDÁN, D., VOJÍŘ, S., ZAVŘEL, M., ŠANCA, O., TACHEZY, R. (2023). Can variants, reinfection, symptoms and test types affect Covid-19 diagnostic performance? A large-scale retrospective study of AG-RDTs during circulation of Delta and Omicron variants, Czechia, December 2021 to February 2022 [online]. *Euro Surveill*, 28(38): 2200938. <https://doi.org/10.2807/1560-7917.ES.2023.28.38.2200938>.

KOREVAAR, D. A., TOUBIANA, J., CHALUMEAU, M., MCINNES, M. D. F., COHEN, J. F. (2021). Evaluating tests for diagnosing Covid-19 in the absence of a reliable reference standard: pitfalls and potential solutions [online]. *Journal of Clinical Epidemiology*, 138: 182–188. <https://doi.org/10.1016/j.jclinepi.2021.07.021>.

KUCHARSKY, Š., WAGENMAKERS, E. (2023). *Correct Conclusions from Fallible Medical Tests: a Tutorial with JASP* [online]. 2023, December 18. <https://doi.org/10.31219/osf.io/jksz6>.

MANI, S., SCHUT, A. (2023). The impact of the Covid-19 pandemic on inequalities in preventive health screenings: Trends and implications for U.S. population health [online]. *Social science & medicine* (1982), 328: 116003. <https://doi.org/10.1016/j.socscimed.2023.116003>.

MOSSMAN, D., BERGER, J. O. (2001). Intervals for posttest probabilities: a comparison of 5 methods [online]. *Med. Decis. Making*, 21(6): 498–507. <https://doi.org/10.1177/0272989X0102100608>.

MORGAN, D. J., PINELES, L., OWCZARZAK, J., MAGDER, L., SCHERER, L., BROWN, J. C., REECE, D. et al. (2021). Accuracy of practitioner estimates of probability of diagnosis before and after testing [online]. *JAMA Internal Medicine*, 181(9): 1174–1181. <https://doi.org/10.1001/jamainternmed.2021.2225>.

PADHYE, N. (2022). Bayesian measurement of diagnostic accuracy of the RT-PCR test for Covid-19. *Metrology*.

POLLATSEK, A., WELL, A.D., TYGERT, M. (1987). Understanding conditional probabilities [online]. *Organizational Behavior and Human Decision Processes*, 40(2): 255–269. <https://doi.org/10.1016/0749-5978(87)90050-5>.

RAIFFA, H., SCHLAIFER, R. (1961) *Applied Statistical Decision Theory*. Boston: Graduate School of Business Administration, Harvard University.

SISAY, A., ABERA, A., DUFERA, B., ENDRIAS, T., TASEW, G. (2022). Diagnostic accuracy of three commercially available one step RT-PCR assays for the detection of SARS-CoV-2 in resource limited settings [online]. *Plos One*. <https://doi.org/10.1371/journal.pone.0262178>.

WANG, H., LIU, Q., HU, J., ZHOU, M., YU, M. Q., LI, K. Y., XU, D., XIAO, Y. et al. (2020). Nasopharyngeal Swabs Are More Sensitive Than Oropharyngeal Swabs for COVID-19 Diagnosis and Monitoring the SARS-CoV-2 Load [online]. *Frontiers in Medicine*, 7: 334. <https://doi.org/10.3389/fmed.2020.00334>.

WAGENMAKERS, E.-J., LOVE, J., MARSMAN, M., JAMIL, T., LY, A., VERHAGEN, J., MOREY, R. D. (2021). Bayesian Inference for Psychology. Part I: Theoretical Advantages and Practical Ramifications [online]. *Psychonomic Bulletin & Review*, 28: 51–71. <https://doi.org/10.3758/s13423-020-01746-2>.

WILLIAMSON, S. F., WILLIAMS, C. J., LENDREM, B. C. (2023). Sample size determination for point-of-care Covid-19 diagnostic tests: a Bayesian approach. *Diagnostic and Prognostic Research*.

ZHANG, Z., DU, X. (2020). Bayesian inference and simulation on true prevalence of the coronavirus disease (Covid-19) in China and its screening diagnostic evaluation [online]. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3569885>.