Report 4: COVID-19: analysis of two serological tests

Marco Colocrese, s301227,
ICT for Health attended in A.Y. 2021/22

December 8th, 2021

1 Introduction

Spreading of COVID-19 (COrona VIrus Disease 19) infection can be reduced with early detection of infected people, so that they can start quarantine as soon as possible. The nasopharyngeal swab test is highly reliable but it requires time and is expensive, serological tests are faster and cheaper, but less reliable. Serological tests find the presence of IgG (Immunoglobulin G) and a high level of this antibody in blood means that the person is or has been affected by COVID-19.

This work reports the results of the analysis of two serological tests, discussing the setting of the thresholds to declare a positive result.

2 Method

A group of 879 people was subjected to 3 tests: one nasopharyngeal swab test and two serological tests (Test 1 and Test 2 in the following), recording the amount of IgG; 17 cases were removed from the dataset due to an uncertain swab test result. The positive swab tests were 71, whereas the negative ones were 791.

Test 1 contained 3 outliers, which were identified using DBSCAN [1] with parameters $epsilon = 8$ and $M = 2$, and then removed (only from Test 1). In these cases the swab test was negative for the three tests.

Swab test result was considered correct, and ROC curve (sensitivity versus false alarm, see Fig. 1) was measured for the two serological tests. The area under ROC was measured equal to 0.937 for Test 1 and 0.928 for Test 2.

For convenience, sensitivity and specificity versus threshold are also plotted in Fig. 2.

The following notation will be used: D means that the patient is really infected, H means that the patient is healthy, T_p means that the test is positive, T_n mean that te test is negative.

The general approach in case of a positive serological test is to check again the person using the nasopharyngeal swab. This makes acceptable a relatively large false positive prob-

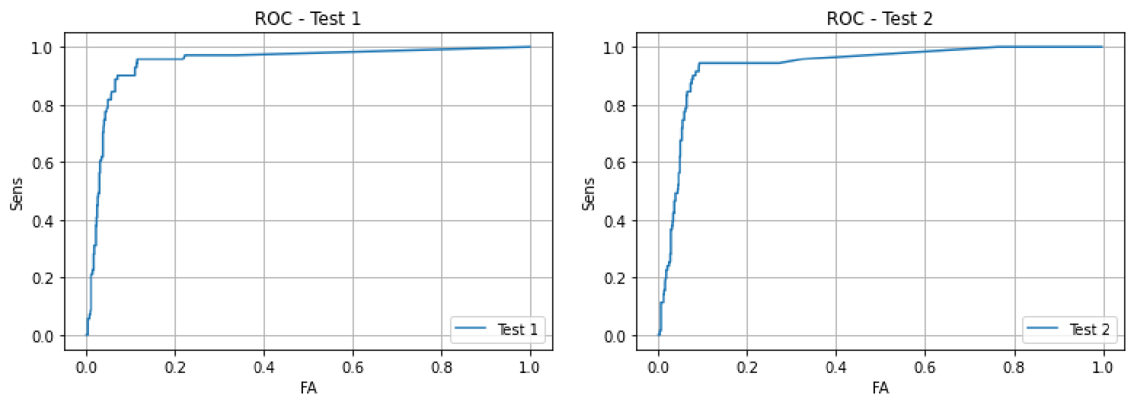

Figure 1: ROC curve for Test 1 (left) and Test 2 (right).

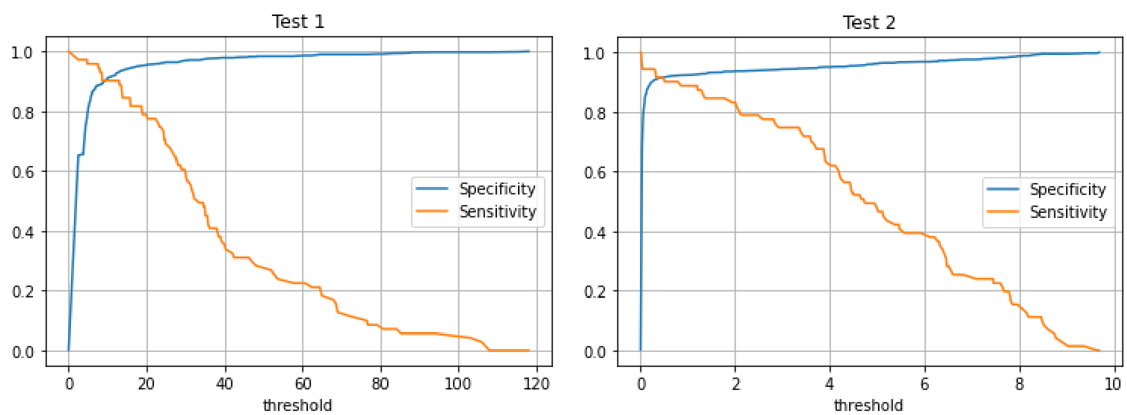

Figure 2: Sensitivity and specificity versus threshold for Test 1 (left) and Test 2 (right).

ability $P(T_p|H)$, with the only drawback that healthy people stay for a couple of days at home, maybe in an anxious state.

What cannot be accepted is instead a large false negative probability: in this case nasopharygeal swab is not tested, and the person can spread the virus to many others. Thus, it is important to have a large sensitivity $P(T_p|D)$ (probability that the test is positive given that the person has the disease), but even more important is the probability $P(D|T_n)$ that the person has the disease given the test is negative. This last probability should be kept as small as possible.

Having assumed COVID-19 prevalence equal to 2 %, Figs. 3 and 4 respectively show versus the threshold:

1. The probability $P(D|T_p)$ that the patient is truly infected given that the test is positive and the probability $P(D|T_n)$ that the patient is infected given that the test is negative.

2. The probability $P(H|T_p)$ that the patient is healthy given that the test is positive and the probability $P(H|T_n)$ that the patient is healthy given that the test is negative.

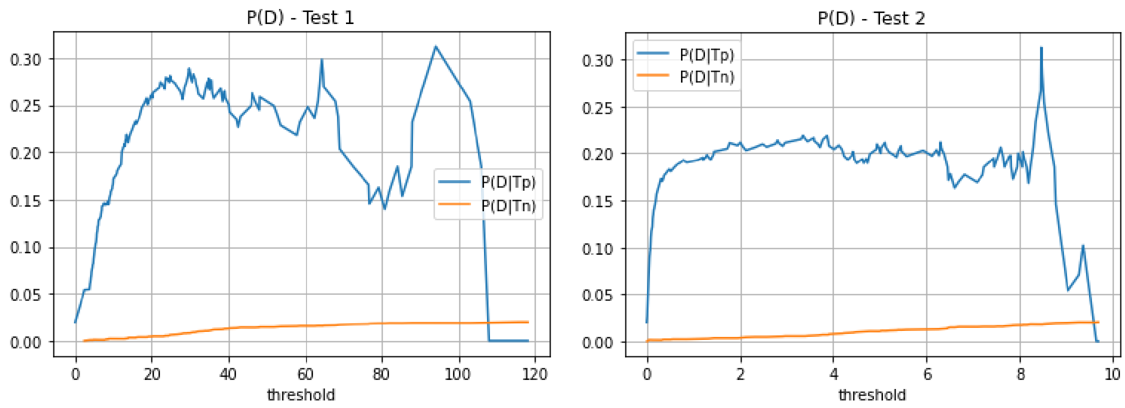

Figure 3: $P(D|T_p)$ and $P(D|T_n)$ versus threshold for Test 1 (left) and Test 2 (right).

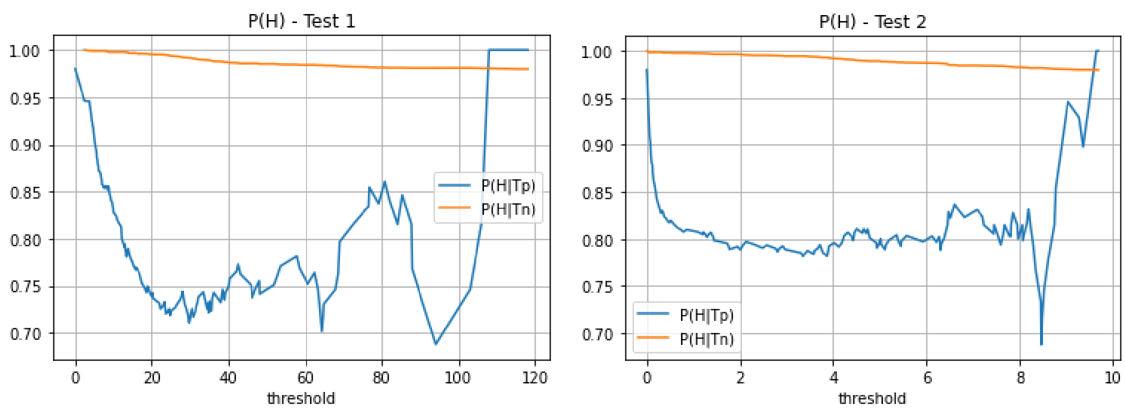

Figure 4: $P(H|T_p)$ and $P(H|T_n)$ versus threshold for Test 1 (left) and Test 2 (right).

3 Choice of the threshold

3.1 Test 1

- Sensitivity and specificity are both equal to 0.901 when the threshold is equal to 9.28, in which case $P(D|T_n) = 0.002$ and $P(D|T_p) = 0.157$.
- Being this sensitivity not large enough, it is convenient to decrease the threshold; we suggest a threshold equal to 7.59 for which:
  - Sensitivity $P(T_p|D) = 0.958$, false negative probability $P(T_n|D) = 0.042$.
  - Specificity $P(T_n|H) = 0.886$, false positive probability $P(T_p|H) = 0.114$.
  - $P(D|T_n) = 0.001$, $P(H|T_n) = 0.999$.
  - $P(D|T_p) = 0.146$, $P(H|T_p) = 0.854$.

3.2 Test 2

- Sensitivity and specificity are both equal to 0.915 when the threshold is equal to 0.44, in which case $P(D|T_n) = 0.002$ and $P(D|T_p) = 0.181$.
- Since this sensitivity cannot be considered sufficient, the lower threshold 0.3 is recommended, for which:
  - Sensitivity $P(T_p|D) = 0.944$, false negative probability $P(T_n|D) = 0.056$.
  - Specificity $P(T_n|H) = 0.908$, false positive probability $P(T_p|H) = 0.092$.
  - $P(D|T_n) = 0.001$, $P(H|T_n) = 0.999$.
  - $P(D|T_p) = 0.173$, $P(H|T_p) = 0.827$.

4 Conclusions

It is well-known that COVID-19 is very contagious, as its transmission is easily made through person-to-person (it is zoonotic, but person-to-person contact resulted the most harmful) contact via respiratory droplets from coughing, sneezing, or talking. One of the major issues related to prevention is the detection of symptoms: a recent systematic review and meta-analysis estimated that the proportion of total infections that are truly asymptomatic ranges from 6% to 41% (pooled estimate of 15%).

Furthermore, a theory by Beldomenico (Int J Infc Dis 2020) sustains that the variability in the global incidence of COVID-19 may be explained by the presence of superspreaders (asymptomatic subjects that can easily spread the virus). Before the availability of the vaccine, there were no pharmacologic preventive interventions, therefore the nonpharmacologic ones assumed a central role in spread restricting. Even with the arrival of the vaccine, these continue to have significant importance. Two of these strategies are quarantining close contacts of infected individuals and the identification and a quick test of suspected cases with subsequent isolation of infected individuals. For this purpose, serological tests resulted essential to give a fast and cheap response, even though less reliable than nasopharyngeal swab. Those tests are recommended to people with COVID-19 symptoms (fever, cough, shortness of breath and breathing difficulties, headache, loss of smell, myalgia, rhinorrhea, taste dysfunction and sore throat), close contacts (within 6 ft), people working in a nursing home or long-term care facility (test regularly) and workers in critical infrastructures and in health care.

The two analyzed tests have given similar results. Overall, the plot of the two ROC curves (Fig. 1) and the values of their area show that Test 1 performs better than the second, being the ideal area under a ROC curve equal to 1.

Observing the results through a more specific lens, the most important parameters, in this case, are sensitivity and the false-negative rate $P(T_n|D)$. It is important to consider the last one to have the near-certainty that a person that had a negative test could not spread the virus, not being actually healthy, and should be close to 0. The sensitivity should be as close as possible to 1. It is not possible to reach this ideal value because it would mean selecting a threshold near zero, which would make the test useless as all the tests would be positive. Plots (Figs. 3 and 4) show that a little threshold enough grater than zero (left side of the plots) gives values of $P(D|T_n)$ enough close to 0 and values of $P(H|T_n)$ enough close to 1 (as it should be), maintaining acceptable values for $P(H|T_p)$ and $P(D|T_p)$.

However, the sensitivity values lead us to prefer test 1. In fact, to obtain a similar value of sensitivity the test 2, the chosen threshold will make the false alarm rate greater than 0.3. The false alarm is not as important as the analyzed values, but a too-large value could make medical doctors and common people give less credibility to a positive test. In this sense, for the chosen thresholds, test 2 has a better behaviour; nevertheless, a value of 0.114 (test 1) can be considered acceptable, considering the related sensibility and $P(T_n|D)$.

While it is fast, cheap and easy to perform, there are several limitations to serological assays, in particular the immunoassays. It is related to the fact that antibodies take several (approximately 3-4) days to be generated after exposure to foreign material, making the test ineffective during this period[2].

In any case, this analysis shows that it is not possible to have certainty, also after the incubation time. Therefore, medical doctors, if there is a high suspicion for COVID-19 (symptomatic individuals), recommend the repetition of the test if it gave a negative result.

References

[1] Ester M., Kriegel H-P., Sander J., Xu X. *A density-based algorithm for discovering clusters in large spatial databases with noise*. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining 1996 (KDD-96), pp. 226-231. AAAI Press.

[2] https://www.frontiersin.org/articles/10.3389/fimmu.2020.00879/full