Estimating the False-Positive Probability from a Recent Covid Study

Michael J. Neely
University of Southern California

Abstract

This short note considers a recent study of the rapid antigen Covid test for 903408 asymptomatic people in Canada during 2021. The study finds that 41.9% of the people who tested positive did not in fact have the disease. While this seems to indicate a large amount of false-positive error, it does not necessarily imply that the rapid antigen test is inaccurate. True figures of merit of the rapid antigen test are the $FalsePositive$ and $FalseNegative$ probabilities. Unfortunately, the study cannot determine empirical estimates of these values because it does not provide ground-truth data for individuals who test negative. However, this note introduces a reasonable assumption on the a-priori probability that a member of the experimental group has the disease. Under this assumption, the empirical $FalsePositive$ probability for this experiment is shown to lie in the interval $[0.000613, 0.000718]$. This means that the empirical measurement for the conditional probability of testing positive, given the individual does not have the disease, is between 0.0613% and 0.0718%. The final section provides a related extra credit assignment for the EE 503 probability class.

I. Introduction

This note discusses the recent article in [1], which can be found at the following link:
`https://jamanetwork.com/journals/jama/fullarticle/2788067`

The article [1] provides data on 903408 rapid antigen tests of Covid in asymptomatic people in Canada during 2021. Of these, 1322 cases tested positive. A subset of these 1322 cases were followed up with more extensive PCR tests that are assumed to act as "ground truth" for determining if an individual truly has the disease. The subset of these positive cases consisted of 1103 people, and of these there were 462 that were determined to not have the disease via the more extensive PCR followup test. From these empirical values one can form the estimates:

$$P[Positive] \approx \frac{1322}{903408} \approx 0.00146 \tag{1}$$

$$P[Have^c|Positive] \approx \frac{462}{1103} \approx 0.419 \tag{2}$$

where $Positive$ is the event that a randomly selected individual tests positive in the rapid antigen test; $Have$ is the event that a randomly selected individual has the disease; $Have^c$ is the *complement* of the event $Have$ (so that $Have^c$ is the event that a randomly selected individual does *not* have the disease); $P[Positive]$ is the probability that the event $Positive$ occurs; $P[Have^c|Positive]$ is the conditional probability that the event $Have$ does not occur, given that the event $Positive$ occurs. Specifically, one can imagine the randomly selected individual as being chosen at a single time, sometime during the year 2021, from the population of all people living in Canada who were represented by this experiment and who were asymptomatic at the time they were chosen. The values in (1)-(2) are probabilities (in the interval $[0, 1]$) and represent: (i) An a-priori chance of 0.146% of testing positive for the disease; (ii) An a-posteriori chance of 41.9% of not having the disease given a positive result on the rapid antigen test.

Unfortunately, data on ground truth PCR testing for individuals who tested *negative* for the rapid antigen test was either not collected or not reported in [1]. This significantly limits the conclusions that can be drawn about the performance of the rapid antigen test. It also significantly limits the ability to estimate the prevalance of the disease within the asymptomatic population considered in [1] at the time the experiment was conducted. This note establishes the (wide) range of performance that can be inferred by the existing data. It also narrows the performance range by making an educated guess about the value of the missing data.

The approximations (1) and (2) assume the sample size is large enough so that the empirical ratios $1322/903408$ and $462/1103$ accurately represent $P[Positive]$ and $P[Have^c|Positive]$, respectively. The approximation (2) also assumes that no additional dependencies were introduced in selecting the 1103 people who participated in the followup PCR testing from the larger pool of 1322 people who tested positive in the rapid antigen test. In particular, it is assumed that the 1103 people were selected uniformly with all selections equally likely. For simplicity of exposition, this short note focuses only on the empirical ratios and does not provide measures of statistical significance such as confidence intervals. This note also ignores complicating factors such as the PCR test itself being an imperfect measure; and issues of timing consistency between the rapid antigen test and the PCR test (a duration of 2 hours can produce different results than a duration of 2 days because the individual's Covid status can change in between tests).

A. *Interpretation and intuition*

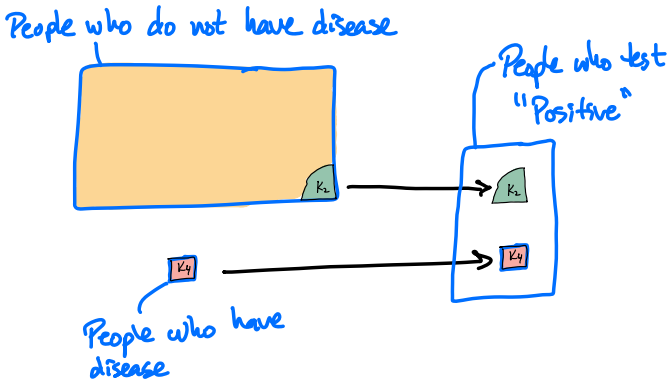

Fig. 1. An example of disease testing where a large group of people who do not have the disease contains a very small portion consisting of K_2 people who test as a false positive; while a very small group of people who have the disease are all perfectly detected as positive. This shows a situation where a highly accurate test has a perfect $FalseNegative$ value of 0, a very low $FalsePositive$ value, but where approximately half of all people who tested positive do not have the disease.

The approximation (2) indicates that there is a 41.9% chance that an individual who tested positive in the rapid antigen test does *not* in fact have the disease.[1] Does this suggest that the rapid antigen test produces an unacceptably large number of false positives? Not necessarily. True figures of merit for the rapid antigen test are the $FalsePositive$ and $FalseNegative$ probabilities, defined as the following conditional probabilities:

$$FalsePositive = P[Positive|Have^c]$$
$$FalseNegative = P[Positive^c|H]$$

Specifically, $FalsePositive$ is the conditional probability that the rapid antigen test gives an erroneous positive result, given that the individual being tested does not have the disease; $FalseNegative$ is the conditional probability that the rapid antigen test gives an erroneous negative result, given that the individual being tested has the disease. In principle, $FalsePositive$ and $FalseNegative$ are fixed values that depend only on the rapid antigen test itself. These values do not vary with the prevalence of the disease within the general asymptomatic population (this prevalence can change over time). In contrast, the reported value $P[Have^c|Positive] \approx 0.419$ depends heavily on the disease prevalence at the time the experiment was conducted.

In any disease testing scenario there is a tradeoff between $FalsePositive$ and $FalseNegative$ values. The performance of the rapid antigen test cannot be ascertained unless both values are considered. For example, a test that always outputs "Negative," regardless of the biological data of the individual being tested, would result in a perfect $FalsePositive$ score of 0 but would have an abysmal $FalseNegative$ score of 1. This type of undesirable

[1]The article [1] provides a detail that seems to attribute the false positives to two particular testing locations and one particular "lot" or "batch" of data. It was unclear how many "lots" of data there were and how many people were in those two locations. This detail is ignored in the current note.

test has no ability to detect the disease but would appear to perform well if $FalsePositive$ were the only objective to be considered.

It is possible for the rapid antigen test to have low values for both $FalsePositive$ and $FalseNegative$ while still producing an expected number of positive cases of 1103 and an expected number of false positive cases of 422 (see an example in Fig. 1). This behavior can arise when the disease prevalence is low, that is, when the a-priori probability of having the disease is a small value (which seems to be the case in the experiment of [1]).

Specifically, let $n = 903408$ be the number of people who participated in the experiment. This group of people can be partitioned into four disjoint subgroups with sizes K_1, K_2, K_3, K_4 each:

- Subgroup 1 contains K_1 people who do not have the disease and test negative in the rapid test.
- Subgroup 2 contains K_2 people who do not have the disease and test positive in the rapid test.
- Subgroup 3 contains K_3 people who have the disease and test negative in the rapid test.
- Subgroup 4 contains K_4 people who have the disease and test positive in the rapid test.

If the values K_1, K_2, K_3, K_4 were known, the following empirical approximations could be used:

$$P[Positive] \approx \frac{K_2 + K_4}{n} \tag{3}$$

$$P[Have^c|Positive] \approx \frac{K_2}{K_2 + K_4} \tag{4}$$

$$FalsePositive \approx \frac{K_2}{K_1 + K_2} \tag{5}$$

$$FalseNegative \approx \frac{K_3}{K_3 + K_4} \tag{6}$$

The right-hand-side of the approximate equalities (3)-(6) are empirical ratios. Can nonnegative integers K_1, K_2, K_3, K_4 be found that sum to n, that yield empirical ratios matching the $P[Positive] \approx 0.00146$ and $P[Have^c|Positive] \approx 0.419$ results from the experiment in (1)-(2), and that yield low values for both the $FalsePositive$ and $FalseNegative$ empirical ratios? Yes: Consider the following *example numbers*:

$$(K_1, K_2, K_3, K_4) = (902086, 553, 0, 769) \tag{7}$$

which yield

$$\frac{K_2 + K_4}{n} = 0.00146$$
$$\frac{K_2}{K_2 + K_4} = 0.418$$
$$\frac{K_2}{K_1 + K_2} = 0.000613$$
$$\frac{K_3}{K_3 + K_4} = 0$$

These example numbers give empirical ratios for $P[Positive]$ and $P[Have^c|Positive]$ that closely match the observed 0.00146 and 0.419 values from the experiment, and yet also have low empirical $FalsePositive$ and $FalseNegative$ ratios of 0.000613 and 0, respectively (where 0.000613 represents a 0.0613% chance of false positive given an individual does not have the disease). No test can have a FalseNegative ratio better than 0, and, arguably, any test that provides a $FalsePositive$ ratio as low as 0.000613 is desirably accurate. Intuitively, the reason why such an accurate test can produce a batch of $K_2 + K_4$ people who test positive, where K_2 and K_4 are on the same order of magnitude, is that the subgroup with K_2 people is formed by taking a very small fraction of the very large pool of people who do not have the disease, while the subgroup with K_4 people is formed by taking a very large fraction of the very small pool of people who have the disease (see Fig. 1).

We call the above ratios *empirical ratios*, and the numbers K_1, K_2, K_3, K_4 *empirical numbers*, because they exist as a particular realization of the experiment and, in principle, they could have been exactly determined if the followup ground truth PCR testing was done on all 903408 participants (regardless of their positive or negative

result in the rapid antigen test). What are the true K_1, K_2, K_3, K_4 values from the experiment? The study establishes the following indirect information:

$$K_1 + K_2 + K_3 + K_4 = n = 903408 \tag{8}$$

$$K_2 + K_4 = 1322 \tag{9}$$

$$\frac{K_2}{K_2 + K_4} \approx \frac{462}{1103} \tag{10}$$

This is not enough information to determine K_1, K_2, K_3, K_4. The example numbers in (7) were chosen to be consistent with (8)-(10). This illustrates that the situation in (7) is *possible*. Therefore, *it is entirely possible that the rapid antigen test provides desirable figures of merit for* $FalsePositive$ *and* $FalseNegative$. However, the example numbers in (7) are not the only ones that are consistent with (8)-(10). The next subsection provides upper and lower bounds on the empirical estimate of $FalsePositive$ based on the information given in [1]. Unfortunately, the upper and lower bounds do not provide insight because they have a huge range. These are essentially the best bounds that can be obtained under the given information, without introducing additional assumptions. Next, a reasonable additional assumption is made to establish a better estimate of the true value of the empirical estimate for $FalsePositive$.

II. Upper and lower bounds

This section provides upper and lower bounds on the figure of merit $FalsePositive$. We have

$$\begin{aligned} FalsePositive &= P[Positive|Have^c] \\ &\overset{(a)}{=} \frac{P[Have^c|Positive]P[Positive]}{P[Have^c]} \\ &\overset{(b)}{=} \frac{\left(\frac{462}{1103}\right)\left(\frac{1322}{n}\right)}{P[Have^c]} \end{aligned} \tag{11}$$

where (a) uses Baye's rule; (b) uses $n = 903408$ with (1) and (2) (for simplicity of exposition we treat the approximations $P[Positive] \approx 1322/903408$ and $P[Have^c|Positive] \approx 462/1103$ and as if they are exact). The value of $P[Have^c]$ is the probability of not having the disease and is:

$$\begin{aligned} P[Have^c] &= P[Have^c|Positive]P[Positive] + P[Have^c|Positive^c]P[Positive^c] \\ &= \left(\frac{462}{1103}\right)\left(\frac{1322}{n}\right) + P[Have^c|Pos^c]\left(\frac{n-1322}{n}\right) \end{aligned} \tag{12}$$

where, for simplicity, the approximations are again treated as being exact. Unfortunately, [1] provides no insight into the value $P[Have^c|Positive^c]$ because people who tested negative for the disease did not have followup PCR testing to reveal ground truth. The only available upper and lower bounds on $P[Have^c|Pos^c]$ are the trivial ones:

$$0 \le P[Have^c|Positive^c] \le 1$$

Substituting the upper and lower bounds of 0 and 1 into (12) gives

$$\left(\frac{462}{1103}\right)\left(\frac{1322}{n}\right) \le P[Have^c] \le \left(\frac{462}{1103}\right)\left(\frac{1322}{n}\right) + \left(\frac{n-1322}{n}\right)$$

Substituting these bounds on $P[Have^c]$ into (11) gives

$$0.000613 \le FalsePositive \le 1 \tag{13}$$

The example numbers in (7) provide a situation when $FalsePositive$ matches the lower bound 0.000613. Alternatively, the (unlikely) case when $P[Have^c|Positive^c] = 0$ provides a situation where the upper bound of 1 is met with equality. Thus, these are the best possible upper and lower bounds. However, the interval $[0.000613, 1]$ is too wide to be of use. It is reasonable to make an educated guess that significantly reduces the width of the interval. Specifically, it is reasonable to assume that $P[Have]$, the probability that a randomly selected asymptomatic person has the disease, satisfies

$$0 \le P[Have] \le 100P[Positive] \tag{14}$$

where $P[Positive] = 1322/n$ is the empirical probability of testing positive in the rapid antigen test. The assumption (14) seems to be mild because it allows the empirical probability of testing positive to differ by as much as two orders of magnitude from the true probability of having the disease. Of course, it should be emphasized that Assumption (14) is only a guess and is not supported by any data in [1]. Assumption (14) implies

$$P[Have^c] = 1 - P[Have] \in \left[1 - 100\left(\frac{1322}{n}\right), 1\right]$$

Substituting these bounds for $P[Have^c]$ into (11) yields that, under the assumption (14):

$$\boxed{0.000613 \leq FalsePositive \leq 0.000718}$$

It should be emphasized that the above upper and lower bounds are actually bounds on the *empirical* false positive value $K_2/(K_1 + K_2)$ based on an incomplete knowledge of K_1 and K_2. A statistical analysis of the confidence associated with using this empirical ratio as an estimate of the true $FalsePositive$ value of the rapid antigen test is not given in this note. Such an analysis would be expected to slightly widen the interval $[0.000613, 0.000718]$ without changing the order of magnitude of the endpoints.

III. Extra Credit EE 503 — Probability

a) Check this note for typos, mistakes, and consistency with [1]. Report on any issues that you might find.

b) Determine upper and lower bounds on the empirical value of $FalseNegative$ when (i) $0 \leq P[Have|Positive^c] \leq 1$; (ii) Bounds on $P[Have|Positive^c]$ are inferred by the assumption (14).

c) Write two paragraphs that describe the meaning of the two sentences in the following excerpt from Section I: *The approximation* (2) *also assumes that no additional dependencies were introduced in selecting the 1103 people who participated in the followup PCR testing from the larger pool of 1322 people who tested positive in the rapid antigen test....For simplicity of exposition, this short note focuses only on the empirical ratios and does not provide measures of statistical significance such as confidence intervals.*

References

[1] Joshua S. Gans, Avi Goldfarb, Ajay K. Agrawal, Sonia Sennik, Janice Stein, and Laura Rosella. False-Positive Results in Rapid Antigen Tests for SARS-CoV-2. *JAMA*, 327(5):485–486, 02 2022.