Characterizing medical test errors

Douglas A. Kerr

Issue 2
June 21, 2026

INTRODUCTION

An important class of medical tests seeks to determine whether or not the patient has a certain "condition". Ordinarily, a test of this type can return one of two results, "positive" (meaning that the patient does have the condition of interest) and "negative" (meaning that the patient does not have the condition of interest).

But these tests are not "infallible". On some occasions, for a patient **not having** the condition, the test will report "positive" (a "false positive"), and/or on some occasions, for a patient **having** the condition, the test will report "negative" (a "false negative").

Various metrics are used to characterize the performance of a test in this regard.

Imagine a certain test (whose "false" report probabilities are known). If the result is "positive", what is the probability that the patient actually **has** the condition? If the result is "negative", what is the probability that the patient actually **does not have** the condition? This is not as simple as it might seem. Again, there are metrics that express those probabilities.

1 DISCLAIMER

I am not a trained medical processional. The material in this article is from a basic mathematical perspective, and reflects my own understanding of the situation. Those in the medical profession may have other outlooks on the matter.

Nothing here should be taken as the basis for any person to make medical decisions based on the results of medical testing.

2 SUBTLETIES

There are various subtleties to this matter that I ignore because they are not really critical to the story to be told here, and "getting into the weeds" can obscure that story.. For example, I simplistically speak in terms of a patient either having a certain condition or not. But it of course may be more subtle than that.

3 TERMINOLOGY AND LANGUAGE

3.1 "Positive" and "negative"

It is customary in this field that when a test of the type of interest here concludes that the patent **does** have the condition tested for, the result is reported as "positive"; when a test concludes that the patent **does not** have the condition tested for, the result is reported as "negative". I will follow that here.

3.2 True and false

Although the type of test of interest here gives a "Boolean" result (two possible values), the terms "true" and "false" here are not used in the Boolean sense but rather to mean, or a test result, "correct" and "incorrect", respectively

4 THE ERROR BEHAVIOR OF A TEST

4.1 Introduction

Almost universally, a medical test for a certain condition cannot be expected to, in every case, give a result that assuredly describes whether or not the patient has that condition.

Three sets of metrics are commonly used to describe the expected "correctness" of a test, as may have been ascertained by testing done by the test manufacturer or other laboratories..

4.2 The false positive and negative rates

4.2.1 *False positive rate (FPR)*

The *false positive rate* (FPR) for a test is the fraction of the tests given to patients **not having** the condition for which the test nevertheless gives a *positive* result, which is *false*.

The FPR for a test giving no false positive errors is 0.

4.2.2 *False negative rate (FNR)*

The *false negative rate* (FPR) for a test is the fraction of the tests given to patients **having** the condition for which the test nevertheless gives a *negative* result, which is *false*.

The FNR for a test giving no false negative errors is 0.

4.3 Sensitivity and specificity

4.3.1 *Sensitivity*

The *sensitivity* of a test is the fraction of the tests on patients **having** the condition that give a *positive* result, which is *true*.

Sensitivity is one minus the *false negative rate* (1-FNR).

The term comes from the notion that this metric tells us how reliably the test will (via a positive result) recognize the presence of the condition. (This is a bit of a stretch from other uses of "sensitivity" in science and engineering.)

The *sensitivity* for a test giving no false negative errors is 1.

4.3.2 *Specificity*

The *specificity* of a test is the fraction of the tests on patients not having the condition that give a *negative* result, which is *true*.

Specificity is one minus the *false positive rate* (1-FPR).

The term comes from the notion that this metric tells us how reliably the test limits the reporting of the presence of the condition (via a positive result) to only when the condition is actually present. (This is a bit of a stretch from other uses of "specificity" in science and engineering.)

The *specificity* for a test giving no false positive errors is 1.

4.4 The predictive values (PPV and NPV)

4.4.1 *Introduction*

These metrics do not just depend on the behavior of the test of interest but are also influenced by the *prevalence* of the condition (that is, the probability that any given patient in the population being considered will actually have the condition of interest).

They actually express the probability the a positive or negative test result for a certain patient reflects the patent's true situation as to the condition of interest.

These metrics are discussed in further detail in Appendix A.

4.4.2 *Positive predictive value (PPV)*

The *positive predictive value* (PPV) of a test is the fraction of all positive results that are true (the patient actually has the condition).

It tells, if a patient has received a positive result on the test for a certain condition, the probability that the patient in fact has that condition.

The PPV for a test giving no false positive errors is 1.

4.4.3 *Negative predictive value (NPV)*

The *negative predictive value* (NPV) of a test is the fraction of all negative results that are true (the patient actually does not have the condition).

It tells, if a patient has received a negative result on the test for a certain condition, the probability that the patient in fact does not have that condition.

The NPV for a test giving no false negative errors is 1.

5 INTERPRETATION OF THE TEST RESULT—AN EXAMPLE

5.1 Introduction

Suppose that for some condition the probability that a given patient in the population of interest has it (its *prevalence*)[1] is 0.1%. Suppose that the *false positive rate* of the test for that condition is 0.5% and that the *false negative rate* is 0.4%.

Suppose now that a certain patient is given the test, and the result is *positive*. What is the probability that the patient actually has the condition?

It is tempting to think that the answer is 99.5% (the *specificity* of the test). But that is not correct (not even close).

5.2 An example

In this example, we assume a population of 100,000 patents, for which the *prevalence* of a certain condition is 0.1 %. Then we would expect 100 patients to **have** the condition and 99,900 to **not have** it.

We now imagine giving all these patents the test of interest.

5.3 Test result "positive"

Of the 100 patients that **do** have the condition, we would expect (100×0.004) [2] to receive a *negative* result (false in this case). This calculates as 0.4, but I will round it to 1 (rather than 0) to make the story easier to follow. Thus the remaining 99 patients received a *positive* result (true in this case).

---

[1] As pertains to the "population" the patient is in.

[2] Calculated values have (mostly(been rounded to the nearest integer since we can only speak of integral numbers of patients.

Of the 99.900 patients that **do not have** the condition, we would expect 500 (99,900 × 0.005) to receive a *positive* result (false in this case).

Thus, for the 100,000 patients, 599 would receive a *positive* result. But of those, only 99 (the "true positives") were for patients that actually **had** the condition.

Therefore, for a patient for whom the test reported *positive*, the probability that the patient actually **did have** the condition is 99/599, or about 16.5%.

If I had not rounded the values to integers (fudging in one case), the result would be exactly the calculated value of the PPV for this situation.

It may be surprising that this value is so small.

5.4 Test result "negative"

Of the 99,900 patients that **do not have** the condition, 99,900 × 0.005 would receive a *positive* result (false in this case). That calculates as 499.5, but I will round it to 499 (rather than 500) to make the story easier to follow. Thus, the remaining 99,401 patients **not having** the condition would receive a *negative* result (true in this case).

Of the 100 patients that **do have** the condition, 100 × 0.004 would receive a *negative* result (false in this case). This calculates as 0.4, but I will round it to 1 (rather than 0) to make the story easier to follow.

So overall, 99,402 patients would have received a *negative* result. But 1 of those was a "false negative", so there were only 99.401 "true negative" results (over the patients who actually **did not have** the condition).

Therefore, for a patient for whom the test reported *negative*, the probability that the patient actually **did not have** the condition is 99,401/99,402, or 99.9999+%.

If I had not rounded the values to integers (fudging in a couple of case), the result would be exactly the calculated value of the NPV for this situation.

It may be surprising that this value is so near to 100%..

6 CLASSICAL STATISTICS

This matter can be nicely characterized mathematically in terms of conditional probabilities, which are of the form, “If A is true (in the Boolean sense), then what is the probability that B will be true?”. To know the probability that B is true, we must have that conditional probability and must know the probability that A is true.

Recasting the matter of errors in medical diagnostic tests into that light is beyond the scope of this article.

7 BAYESIAN INFERENCE

This matter is often cited as a easily-understood example of Bayesian inference. This is an approach to statistics based on the work of Thomas Bayes (1701-1761).

Simplistically, in Bayesian inference, we reckon the probability of a certain event happening from the combination of:

a. What we know or assume about the probability of the event occurring before taking into account certain new information.

b. New information

In the matter that is the topic of this article, the “event” is whether a certain patient has a certain condition. Item a is the known (or estimated) *prevalence* of the condition (0.1% in the example), and item b is, as to a certain patient, the result of a test for the condition.

So before we give a person a test for that condition, we can say, “The probability that the patient has this condition is 0.1%”.

But then we give the patient a test for the condition (whose “error” statistics we know, and are as in the earlier example). The test result is positive.

What we can now say is, “The probability that the patient has this condition is a little over 16%.”

Further recasting the matter of errors in medical diagnostic tests into that framework of Bayesian inference is beyond the scope of this article.

8 ACKNOWLEDGEMENT

Thanks to my friend and colleague Asher D. Kelman, PhD, MD (ret) for initially bringing this matter to my attention and illuminating it with clear numerical examples..

-#-

Appendix A
PPV and NPV

A.1 INTRODUCTION

The metrics PPV (positive predictive value) and NPV (negative predictive value) are very useful as they describeg the probability that the result of a diagnostic test actually tells whether or not the patient has the condition tested for.

A.2 THE POSITIVE PREDICTIVE VALUE (PPV)

The *positive predictive value* (PPV) for a test situation is the probability that a positive test result is **true**—that the patient receiving a positive test result indeed **does have** the condition of interest.

It can be defined thus:

$$PPV = \frac{TP}{P} \qquad (1)$$

where TP represents the number of true positive results and P represents the total number of positive results, true or false, both over the same large number of tests.

But we get further insight into the situation from this alternate formula for the PPV:

$$PPV = \frac{SENS \times PREV}{SENS \times PREV + (1 - SPEC) \times (1 - PREV)} \qquad (2)$$

where SENS is the *sensitivity* of the test, SPEC is the *specificity* of the test, and PREV is the *prevalence* of the condition (the fraction of the population that has the condition).

So we see that both the properties of the test (its *sensitivity* and *specificity*) and the *prevalence* of the condition itself figure into the PPV.

The numerator of this equation is the expected number of true positive test results (as a fraction of the population). This quantity is also the first term of the denominator.

The second term of the denominator is the expected number of false positive test results (again as a fraction of the population).

Thus the denominator is the total number of positive test results, true or false (again as a fraction of the population).

Thus the entire fraction—the PPV—is the probability that a positive test result is **true**—that the patient receiving a positive test result indeed **does have** the condition of interest.

A.3 THE NEGATIVE PREDICTIVE VALUE (NPV)

The *negative predictive value* (NPV) for a test situation is the probability that a negative test result is **true**—that the patient receiving a negative test result indeed **does not have** the condition of interest.

It can be defined thus:

$$NPV = \frac{TN}{N} \quad (3)$$

where TN represents the number of true negative results and N represents the total number of negative results, true or false, both over the same large number of tests.

But we get further insight into the situation from this alternate formula for the NPV:

$$NPV = \frac{SPEC \times (1 - PREV)}{SPEC \times (1 - PREV) + (1 - SENS) \times PREV} \quad (4)$$

where the variables are as before.

So we see that both the properties of the test (its *sensitivity* and *specificity*) and the *prevalence* of the condition itself figure into the NPV.

The numerator of this equation is the expected number of true negative test results (as a fraction of the population). This quantity is also the first term of the denominator.

The second term of the denominator is the expected number of false negative test results (again as a fraction of the population).

Thus the denominator is the total number of negative test results, true or false, again as a fraction of the population.

Thus the entire fraction—the NPV—is the probability that a negative test result is **true**—that the patient receiving a negative test result indeed **does not have** the condition of interest.

-#-