Josh Hill
11-20-06

Most doctors, when diagnosing a disease, depend on the false-positive rate for the given test. However, according to Bayes' theorem, the prevalence of that disease in the population must be taken into account. This can result in unnecessary surgery or treatment for a person who actually does not have the disease. Furthermore, many patients may also be confused by the numerical accuracy of a test, especially the false-positive and false-negative rates.

For example, say a given test for a disease is 95% accurate, which means it has a 5% false positive rate and 5% false negative rate. If a patient tests positive, one might be inclined to believe that there is a very high chance of that person having the disease. Bayes' Theorem says:

$$\Pr(D|P) = \frac{\Pr(P|D) * \Pr(D)}{\Pr(P)}$$

Pr(P) is the probability of testing positive, which is the probability of testing positive given the person has the disease plus the probability of testing positive given the person does not have the disease. This expands the equation to:

$$\Pr(D|P) = \frac{\Pr(P|D) * \Pr(D)}{[\Pr(P|D) * \Pr(D)] + [\Pr(P|{\sim}D) * Pr({\sim}D)]}$$

Pr(P|D) is the probability of testing positive while having the disease. The false negative rate, 0.05, is the number of people testing negative while having the disease. All those tested who have the disease and did not test negative must have tested positive, meaning that Pr(P|D) is (1-false negative rate)=(1-0.05). Pr(P|~D) is the probability of testing positive while not having the disease. This is the false positive rate which is 0.05. If only 15 out of every 1000000 people have the disease, Pr(D) is then 15/1000000. Pr(~D) is the percentage of people without the disease, which is 1-Pr(D). Putting these values into Bayes' Theorem gives:

$$\Pr(\mathrm{D}|\mathrm{P}) = \frac{(1-.05) * \frac{15}{1000000}}{(1-.05) * \frac{15}{1000000} + .05 * \left(1 - \frac{15}{1000000}\right)} = 0.000285$$

If a person tests positive in this test for this disease, there is still only a 0.0285% chance they have the disease, not 95% as one would normally assume.

Let's look at a real world test, such as HIV tests. Current HIV tests in the United States have a false positive rate of between 0.0004%-0.0007% and a false negative rate of 0.003% (http://en.wikipedia.org/wiki/HIV_test). Pr(P|D) is the probability that the individual is diagnosed with HIV given that they are indeed infected. Since the false negative rate is the probability of testing negative when actually infected, 1-(false negative rate) is the probability of having a positive test and being infected, which is (1-.00003). The prevalence of HIV in the United States is 0.6% (https://cia.gov/cia//publications/factbook/geos/us.html). Therefore, Pr(D), or the probability of having HIV, is 0.006. Pr(~D) is the probability of not being infected, which is 1-(probability of being infected), which is (1-0.006). Pr(P|~D) is the probability of being diagnosed given the individual does not have the disease, which is the false positive rate. Disregarding that some individuals are at higher risk for HIV than others depending on their behavior, let's calculate the likelihood of an individual having HIV if they have a positive test assuming the higher false positive rate:

$$\Pr(D|P) = \frac{\Pr(P|D) * \Pr\ (D)}{[\Pr(P|D) * \Pr(D)] + \ [\Pr(P|{\sim}D) * Pr({\sim}D)]}$$

$$\Pr(\mathrm{D}|\mathrm{P}) = \frac{(1-.00003) * 0.006}{(1-.00003) * 0.006 + \ .000007 * (1-.006))} = 0.9991$$

Based on this data, if a random individual tests positive for an HIV test, there is a 99.91% chance that they actually do have HIV. If I was performing this test on a patient, and they had a positive test result, I would tell them that they almost certainly have the disease. There are some cases where false positives are known to occur, such as in patients given experiment HIV vaccinations. However, most people have not had one of these trial vaccinations. Further inquiry may be made to see how at risk an individual is, since there is the very slim, 0.09%, chance that the test was incorrect. For example, if the person has not had unprotected sex, used any dirty needles, or done anything making it possible for that person to contract HIV, the likelihood of him/her having HIV decreases dramatically. This can cause the test to be "less accurate" as in the example case given earlier. If the probability of the individual having the disease is very small, then Pr(D) decreases greatly and, as seen in the first example, makes the probability of that individual not being infected, despite testing positive, very small. This is regardless of whether the test has a very low false positive rate.

Another diagnostic test that is very common is the rapid test for Streptococcus, or strep throat. This test is often given in doctors' offices to patients with sore throats. The rapid test has a false positive rate of 5% and a false negative rate of between 10% and 30% (http://webseitz.fluxent.com/wiki/StrepThroat). The prevalence of strep in those individuals who have a sore throat is 15% (http://webseitz.fluxent.com/wiki/StrepThroat). This means that Pr(P|D)=1-(false positive rate)=(1-(0.2±0.1)). Pr(D) is 0.15, since only people who have sore throats would be getting the test. Pr(P|~D) is the false positive rate which is 0.05. Let's look at Bayes' Theorum with these values:

$$\Pr(D|P) = \frac{\Pr(P|D) * \Pr(D)}{[\Pr(P|D) * \Pr(D)] + [\Pr(P|\sim D) * Pr(\sim D)]}$$

$$\Pr(\mathrm{D}|\mathrm{P}) = \frac{(1-0.1)*0.15}{(1-0.1)*0.15+\ 0.05*(1-.15))} =\ 0.761$$

$$\Pr(\mathrm{D}|\mathrm{P}) = \frac{(1-0.3)*0.15}{(1-0.3)*0.15+\ 0.05*(1-.15))} =\ 0.712$$

Therefore, if the rapid test is positive, there is between a 71.2% and 76.1% chance that that individual has strep. In this case, since the treatment for strep is just antibiotics, I would probably tell the patient that they most likely have strep and give them the medication. There would be no negative side effects in the patient from getting treatment in this case if he/she did not have strep. Recently, however, there have been problems with the over-use of antibiotics, causing bacteria to become resistant. This is a major problem in hospitals, and since there is still a good chance that the test is wrong, perhaps a culture should be taken and sent away for a more accurate test before giving out the antibiotics. This test, given the high probability of the individual having strep, Pr(D), is not very accurate at all.

With a strep test, it is common for a patient with a negative test to have a culture taken and sent away. Using Bayes' theorem, the reason is clear:

$$\Pr(D|{\sim}P) = \frac{\Pr({\sim}P|D)*\Pr(D)}{[\Pr({\sim}P|D)*\Pr(D)]+\ [\Pr({\sim}P|{\sim}D)*Pr({\sim}D)]}$$

Pr(~P|D) is the probability of having the disease but having a negative test, or the false negative rate which is between 0.1 and 0.3. Pr(~P|~D) is the probability of not having the disease and testing negative, which is 1-(number of people testing positive and not having the disease)=1-(false positive rate)=(1-0.05). Pr(~D) is the probability of not having the disease, which in this case is 1-0.15. Putting these values into Bayes' Theorem gives:

$$\Pr(D|{\sim}P) = \frac{0.1 * 0.15}{0.1 * 0.15 + (1 - 0.05) * (1 - .15))} = 0.0182$$

$$\Pr(D|{\sim}P) = \frac{0.3 * 0.15}{0.3 * 0.15 + (1 - 0.05) * (1 - .15))} = 0.0529$$

Thus, between 1.82% and 5.29% of those tested actually have strep throat but test negative. This is a relatively high percentage of people being misdiagnosed by this test, which affects on average one out of every 28 patients. This accounts for why those with negative rapid tests have a culture taken to verify the results of the test.

Based on the calculations in these examples, it is clear that the false positive and false negative rates of a test are not to be used for determining the diagnosis of the patient. Rather, that individual's chance of having the disease must be taken into account to get the true accuracy of the test.