Why neuroimaging can't diagnose autism

13

Robert T. Thibault, Lauren Dahl and Amir Raz

Wouldn't it be convenient if you could scan your toddler's brain to test for autism? Or, if you're feeling stressed out and down, to neuroimage your own head and see if you're just going through a phase or whether you have clinical anxiety or depression? What about scanning for schizophrenia, bipolar disorder, obsessive compulsive disorder, or any other mental condition for that matter? If we had this ability, the implications would be widespread and the benefits tangible.

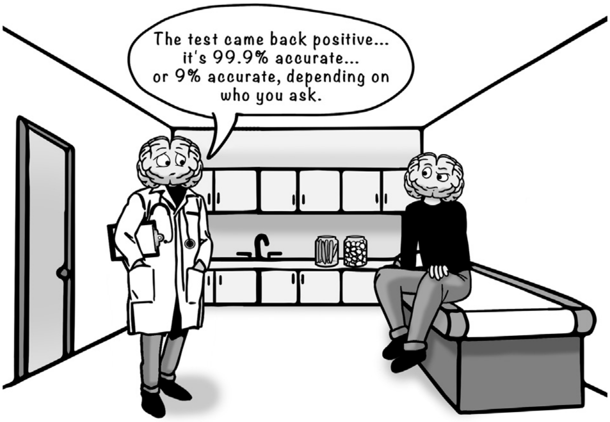
Unfortunately, this option just isn't available yet. Some media reports, however, suggest that the day of diagnostic neuroimaging may be right around the corner. These promises often stem from studies where researchers successfully distinguish between individuals with and without mental conditions: for example, identifying brain scans from people with conditions such as autism spectrum disorder (ASD), major depressive disorder, and schizophrenia with 80%–90% accuracy [1].

With this level of precision, is it only a matter of time before widespread adoption of diagnostic neuroimaging? Not so fast. A simple, but often overlooked, nuance stands in our way. And that nuance lies in the word *accuracy*. Accuracy has both a vague usage in common speech and a very specific meaning in medical diagnoses. Let's use a real life example to illustrate the importance of the different uses of this term [2].

Not too long ago, the physicist Leonard Mlodinow (who notably wrote *A Briefer History of Time* with Stephen Hawking) applied for life insurance and had his blood drawn. As life insurance companies do, they wanted to ensure Mlodinow didn't have any life-threatening illness before insuring him. Surprisingly, the company denied his application for life insurance. Mlodinow's doctor confirmed that he tested positive for HIV and that the test was 99.9% accurate. And yet, Mlodinow remained calm, knowing that this "accuracy" meant he had only a 9% chance of truly being HIV positive.

Wait a second... The doctor says the test is 99.9% accurate and Mlodinow interprets this as a 9% chance of having HIV. What's going on? Fortunately, we don't

Casting Light on the Dark Side of Brain Imaging. DOI: https://doi.org/10.1016/B978-0-12-816179-1.00012-8
© 2019 Elsevier Inc. All rights reserved.



need to be brilliant theoretical physicists of the Mlodinow caliber to understand why he didn't flinch at the test results. The discrepancy between the 99.9% and the 9% manifests because few people in his demographic actually have HIV—the prevalence, or "base rate" is low, about 1 in 10,000, or 0.01%. To use technical terms, the HIV test has a "specificity" of 99.9% (the likelihood that the test appears negative given that the person doesn't have HIV) and a "positive predictive value" of about 9% (the likelihood that the person has HIV given that the test is positive). To better understand this issue, let's look at Table 13.1. It illustrates the outcome of HIV tests on 1,000,000 people of Mlodinow's demographic. While we've included a complete table for those who want a deeper understanding of the issue, we highlighted in yellow the important terms such as base rate, specificity, positive predictive value, and their associated percentages.

As you can see, results from this type of medical test require a bit of interpretation. Both Mlodinow and his doctor were correct; they were simply reporting different statistics: Mlodinow focused on the positive predictive value; his doctor on specificity. In this scenario the positive predictive value constitutes the much more interesting percentage. Once Mlodinow receives a positive test result, he knows the

Table 13.1 Diagnostic chart for a population of 1,000,000 who were HIV tested

Total (1,000,000)	*People with HIV* (100)	*People without HIV* (999,900)	**Base rate** $\frac{100}{1{,}000{,}000} = 0.01\%$
Tested positive (1,095)	True positive (95)	False positive (1000)	**Positive predictive value** $\frac{95}{1095} = 9\%$
Tested negative (998,905)	False negative (5)	True negative (998,900)	Negative predictive value $\frac{998{,}900}{998{,}905} > 99.99\%$
	Sensitivity $\frac{95}{100} = 95\%$	**Specificity** $\frac{998{,}900}{999{,}900} = 99.9\%$	

These tests are commonly considered 99.9% accurate (specificity); meanwhile, the 9% positive predictive value represents a more interesting statistic for those who test positive.

result can only be a true positive or a false positive. Looking at specificity tells him little; looking at the positive predictive value tells him how likely it is that he actually has HIV.[1]

This example almost touches on the true technical diagnostic definition of accuracy: the total number of correctly classified cases divided by the total number of cases (in our example, 95 + 998,900 divided by 1,000,000). For conditions with low base rates the percentage associated with accuracy is often very close to that indicated by specificity. In our example, accuracy is a hair less than 99.9%. In different contexts the word "accuracy" could loosely indicate specificity, the positive predictive value, sensitivity (the likelihood that a test is positive given that the person has the condition) or, the correct technical definition. Now that we've hashed out the statistical concept, let's turn back to diagnostic brain imaging.

By looking at brain scans, researchers can distinguish between individuals with and without ASD with up to 90% accuracy. But does this mean I can bring my toddler to a brain imaging clinic and find out whether she has ASD? As for the HIV example, we need to first determine the base rate for ASD (which is somewhere between 1% and 2%) [3,4]. Now we can flesh out the same Table 13.1 and work backward to calculate the value we are interested in—the positive predictive value. We'll use the results from a widely reported diagnostic neuroimaging study on ASD [5], which has previously been dissected in the *Guardian* [6], for our Table 13.2.

The positive predictive value is about 4%. That means, if we use this test (which legitimately has greater than 80% accuracy in technical terms) for every child correctly diagnosed with ASD, another 24 will be incorrectly diagnosed with ASD. To have a very good positive predictive value, we would need to either look at a condition with a much higher base rate or have very high accuracy (or both). However, given the behavioral and neurological heterogeneity of most psychiatric conditions (outlined in Chapter 3), nearing 100% accuracy may present an intractable task.

As you will appreciate next, whether or not we should use diagnostic neuroimaging depends on much more than the positive predictive value alone. Imagine a scenario where an initial brain scan could identify the development of a tumor with a positive predictive value of 4%. If the 96 out of the 100 people who are incorrectly diagnosed with a brain tumor could immediately go through another test to confirm whether the initial scan was a true positive, little harm would be done. If the remaining four people correctly diagnosed could now receive a treatment to remove the tumor before it causes any serious issues, large benefits would ensue. In this scenario a positive predictive value of 4% could be sufficient to recommend the test. In the case of ASD, individuals incorrectly diagnosed with ASD (or rather their parents) may suffer considerable stress until the child is old enough to test for ASD behaviorally. There's no clear route of further action to take for those correctly diagnosed. Thus in the case of ASD the harms of using a test with a low positive predictive value seem to outweigh the benefits.

[1] In case you're wondering, further tests confirmed that Mlodinow did not have HIV.

Table 13.2 Key diagnostic values for neuroimaging ASD

Total (1000)	*People with ASD* (10)	*People without ASD* (990)	**Base rate** $\frac{10}{1000} = 1\%$
Tested positive (207)	True positive (9)	False positive (198)	**Positive predictive value** $\frac{9}{207} = 4\%$
Tested negative (793)	False negative (1)	True negative (792)	Negative predictive value $\frac{792}{793} = 99.9\%$
	Sensitivity $\frac{9}{10} = 90\%$	**Specificity** $\frac{792}{990} = 80\%$	

While researchers and clinicians continue to conduct studies to improve diagnostic neuroimaging toward a clinically useful state, some practitioners claim that they can already provide this service (see Chapter 22). These declarations have been rejected by major clinical societies and research experts, but these rebukes haven't stopped such practitioners from becoming wealthy at the expense of gullible clients. While diagnostic neuroimaging for psychiatric conditions may well find a place in the future of medicine, it likely won't look anything like today's private clinics. Unfortunately, for these private practitioners, terms such as sensitivity, specificity, and positive predictive value are of little concern.

When discussing the results of any diagnostic test, it can be helpful to think not only of the percentages, but more importantly, precisely what those percentages mean. It could just save you some distress at your next visit to the doctor.

Additional readings

A breezy academic article on diagnostic probabilities: Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. BMJ 2003;327(7417):741–4.

A well written and popular style book on randomness and statistics: Mlodinow L. The drunkard's walk: how randomness rules our lives. Vintage; 2009.