Nicholas Binney
BVet Med MSc & DIC PhD MRCVS

Title:

Meno's Paradox and Medicine

The measurement of diagnostic accuracy is an important aspect of the evaluation of diagnostic tests. Typically, this is done by assessing the performance of the test under evaluation (the index test) against the performance of another test that is already trusted to distinguish between patients with and without the target condition (the gold or reference standard). How, then, can the accuracy of a test be assessed if there is not agreed upon gold standard? How is the accuracy of the gold standard, as the most accurate test of all, to be assessed? Here I draw attention to an important problem in the epistemology of diagnosis, and discuss how it might be addressed.

Sometimes, medical researchers try to discover the set of observations that are most accurate of all by directly inspecting diseased and non-diseased patients to identify their distinguishing characteristics. This method is perhaps intuitively appealing, as it seems a straightforward empirical way of discovering how to identify diseased patients. It amounts to trying to correlate the results of diagnostic tests with disease status, and thus seems to conform to the principles of evidence based medicine, which prizes evidence of correlation above all other considerations. I present four examples of researchers who try to produce definitive diagnostic criteria by directly inspecting diseased and not diseased patients, drawn from the medical literature on the diagnosis of eosinophilia-myalgia syndrome, fibromyalgia, infantile rickets, and abusive head trauma. Despite this method's intuitive appeal I will argue that it is impossible to carry out, and that the researchers I refer to here consequently use circular arguments to support their views. Before researchers can inspect these patients to discover distinguishing characteristics, they must be able to distinguish diseased and not-diseased patients; and they do not know how to do this, because this is what they are trying to discover. I suspect the intuitive appeal of directly inspecting patients makes this difficult to appreciate. To counter this difficulty, I present this problem as a manifestation of 'Meno's paradox', which was described by Plato in classical antiquity, and of 'the problem of nomic measurement', described by Hasok Chang more recently in philosophical literature on the measurement of temperature. Considering these philosophical problems may help researchers address the methodological issues they face when evaluating diagnostic tests.

These issues raise the question of how the gold standard diagnostic test should be justified. Some medical practitioners argue that the 'diagnostic accuracy paradigm' should be abandoned altogether. Instead of searching for diagnostic practices that accurately select the correct group of patients in some deep ontological sense, such researchers advocate searching for diagnostic practices that can be used to improve patient outcomes. This approach, however, is not unproblematic. These researchers recognise that if a group of patients selected using a diagnostic test does not respond well to treatment, then focusing solely on outcomes will not determine whether it is the diagnostic test or the treatment, or both, that is inadequate. Other strategies for evaluating diagnostic tests, in addition to correlating test results with other test results or patient outcomes, need to be developed. One such alternative strategy may be to make use of different methods of inference. Instead of focusing on finding correlations using *induction*, perhaps *abduction* might be a useful method of inference. Indeed, the value of many of the diagnostic practices used to diagnose the diseases I discuss here were originally justified using abductive arguments, not inductive ones. I argue that drawing researchers' attention to the actual and potential role of abductive arguments in the evaluation of diagnostic practices would be of benefit to medical practice.