The modern scientific physician: 4. The useful property of a diagnostic

Olli S. Miettinen

In the process of diagnosis, facts about the manifestations of the underlying, unknown illness are secured, in addition to facts that bear on the person's propensity to have various possible illnesses at the time. The manifestational facts emerge as the result of applying *diagnostics* — from symptom-eliciting questions through elements of physical examination to laboratory tests. ("Diagnostic" as a noun referring to a diagnosis-oriented procedure may still be a bit of a neologism; cf. "preventive"/"prophylactic" and "therapeutic" as nouns.)

For each diagnostic, the diagnostician is supposed to be clear on its relevant 'properties,' if not as a basis for the decision to pursue the corresponding fact(s), then at least for rational interpretation of the fact(s) insofar as the diagnostic indeed is used.[1]

Thoughtful as the scientific diagnostician is, (s)he is not content merely to know the putatively relevant 'properties' of the diagnostics (s)he might use; as a preliminary to this, (s)he contemplates those 'property' concepts per se, critically, so as to decide whether to adopt them in the first place. And mere passing reflection on these 'properties' makes him/her realize, for a start, that the terms "sensitivity" and "specificity" are *misnomers* for what the terms — equally misnomers — "true-positive rate" and "true-negative rate" respectively stand for. However high a 'true-positive rate' — 'sensitivity' — may be, this rate does not mean that the test 'senses' the presence of the illness at issue, as the rate might be just as high in the absence of the illness. And as for 'specificity,' the thoughtful, scholarly diagnostician is, linguistically, seriously challenged to say that whereas, in the diagnosis of whatever particular illness, the result from an arbitrarily chosen diagnostic is highly likely to be negative in the absence of that illness, therefore even an arbitrarily chosen diagnostic thus is prone to be highly "specific" for the particular illness at issue! Properly, as is obvious to a scholar, it is the *result* of a diagnostic's application that can be more or less peculiar — specific — for the presence, or absence, of a particular illness. Finally, "true-positive test result" properly means, simply, test result that is truly positive — quite irrespective of whether the illness at issue is present. The proper meaning of "true-negative test result" is equally obvious.

Linguistic anomalies like these make the thoughtful diagnostician wonder whether the associated concepts themselves are malformed. As a little exercise in 'concept analysis,' (s)he might focus on the 'sensitivity' of, say, a particular question about hemoptysis, or a particular radiographic test, in the diagnosis of lung cancer. In so doing, (s)he would realize that the 'true-positive rate' of each of these diagnostics is highly dependent on how advanced the illness is, in terms of tumour size for example. (S)he thus would realize that 'sensitivity,' with its inherent lack of specificity to any particular subtype of the illness, is a seriously *malformed concept.* And equally malformed is the concept of 'true-negative rate' or 'specificity,' given the lack of specificity in the condition again, the mere absence of the illness at issue.

(S)he then might note that "predictive value positive" and its negative counterpart also are misnomers in that prediction is not what diagnosis is about, and that the corresponding 'properties' involve, again, the conceptual malformation of nonspecificity of the condition — notably in the usual case in which a whole range, or set, of possible results from the diagnostic is classified, simply, as positive or negative. Probability/prevalence of lung cancer conditional on a solitary focal opacity noted on a radiograph but not otherwise specified is a concept just as malformed as the probability of such a positive test result conditional on lung cancer not otherwise specified; and again, equally malformed is the concept of the probability of no lung cancer in the absence of such an opacity, so long as this "absence" remains incompletely specified (notably as to the actual number of focal opacities).

Upon these frustrating, deconstructive realizations (s)he may come to a more fundamental, constructive one: different from the tradition in radiology, which tradition itself is odd, diagnostics in general are *not used in isolation*. Instead, they are applied sequentially. Therefore, there is no justifiable interest in those putative, malformed 'properties' of diagnostics: what truly matters is a diagnostic's performance when used as an add-on to the ones already deployed, its *contextual performance* in this sense.

Toward the decision about using a diagnostic, the diagnostician wonders how much the diagnostic probability might change when the result from this diagnostic also is included in the diagnostic profile. This means that the decision-related useful property of a diagnostic is its discriminatingness — *informativeness* — as to the presence/absence of the illness in this contextual, *marginal* probability-changing sense. On an elderly alcoholic with definite jaundice, the determination of serum bilirubin level is rather redundant and thus not particularly informative about the presence/absence of hepatic cirrhosis, but on this person's equally alcoholic but not clearly jaundiced twin the test is quite informative.

Orientational *measures* of a diagnostic's informativeness naturally are the maximum and minimum of the possibly resultant diagnostic post-test probabilities, while complete characterization of it involves the probabilities with which each of the possible post-test probabilities, or ranges of these, will result from the diagnostic's application; and, naturally, those measures are meaningful only when suitably specific to the context of the diagnostic's potential application, not in terms of the pre-test probability per se but the particulars of the profile leading to this.

Once the result is available, it enters the diagnostic profile as an add-on; and this 'updated' profile is, again, *interpreted as a whole*: upon having asked the question about the prevalence of the illness at issue in instances like this in general, the answer constitutes the diagnostic probability at this new stage of the diagnostic process. *No properties* of the component diagnostics, not even their once-marginal degrees of informativeness, are involved in the interpretation.

As an example of studying the putatively relevant yet mere pseudo-properties, and then the genuinely useful property, of a diagnostic, the work on the ventilation–perfusion test in the diagnosis of pulmonary embolism is illustrative. The famous Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED),[2] true to the still-modern culture in radiology, failed to address the test with specificity to the contexts of the patients' particular, varying pre-test profiles. Thus the artificially singular context of the testing, serving purely radiologic diagnosis, got to be that of an unspecified mixture of profiles suggestive of pulmonary embolism, present within the last 24 hours and prompting referral for the ventilation–perfusion test. The investigators defined, *a priori*, the test's possible results in terms of several unidimensional, ordinal categories such as that of "high probability" of pulmonary embolism, based on 7 descriptors of the image. Then, with various cut-off points distinguishing between the positive and negative result on that merely ordinal scale of *a priori* probabilities, they focussed on the corresponding 'properties' of the test in terms of "sensitivity" and "specificity," as though these were conceptually meaningful and also useful in practice.

A subsequent re-analysis of the PIOPED data was done by others.[3] They used the 7 descriptors of the image as diagnostic indicators. They fitted to the data a logistic-regression model, expressing the prevalence of pulmonary embolism as a joint function of those descriptors of the image. As for the informativeness of the test, then, these latter investigators reported that the resulting probability estimates for the presence of pulmonary embolism ranged from 12% to 90% (depending on the facts concerning the 7 descriptors of the image), with the 25th, 50th and 75th centiles of the distribution equal to 15%, 17% and 67%, respectively.

A diagnostic's known degree of marginal informativeness enables the diagnostician to judge its usefulness as an add-on to the previous ones, before actually deploying it. If it is deployed, this is done for the sole purpose of 'updating' the information and, thereby, the diagnostic probability. Fundamentally different from an intervention, a diagnostic thus is not invoked with the idea that it, in itself, will change the course of health for the better. Thus, resolutely rejected must be the now-eminent idea in radiology[4,5] that the use of a diagnostic constitutes an intervention and thereby has effectiveness (in improving health) as its useful property. Roentgen's discovery of x-rays (in physics) led to the development of an informative radiologic diagnostic for pulmonary tuberculosis, but the development of an effective therapeutic intervention had a very different foundation: the biologic discovery of streptomycin by Waksman and his co-workers.

As for the diagnosis-oriented theoretical writings in radiology — and secondarily in 'clinical epidemiology' — again, "Read not to contradict, nor to believe, but to weigh and consider."[6] The same applies, of course, to the writing here.

Dr. Miettinen is with the Department of Epidemiology and Biostatistics and the Department of Medicine, Faculty of Medicine, McGill University, Montreal, Que., and also with the Department of Medicine, Weill Medical College, Cornell University, New York, NY.

Competing interests: None declared.

Acknowledgement: On the 'properties' of diagnostics I have had extensive, like-minded discussions with two internist colleagues, Drs. J. Jaime Caro and Kenneth M. Flegel, and the presentation here draws, in part, from those discussions.

References

1. Goldman L. Quantitative aspects of clinical reasoning. In: Fanci AS, Braunwald E, Isselbacher KJ, Wilson JD, Martin JB, Kasper DL, et al., editors. *Harrison's. Principles of internal medicine*. 14th ed. vol. 1. New York: McGraw-Hill; 1998.
2. The PIOPED Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism: results of the Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED). *JAMA* 1990;263:2753-9.
3. Miettinen OS, Henschke CI, Yankelevitz DF. Evaluation of diagnostic imaging test: diagnostic probability estimation. *J Clin Epidemiol* 1998;51:1293-8.
4. Fryback DJ, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88-94.
5. Hillman BJ. Outcomes research and cost-effectiveness analysis for diagnostic imaging [editorial]. *Radiology* 1994;193:307-10.
6. Bacon F. Essays, 1597. In: Vickers B, editor. *Francis Bacon. The essays or counsels civil and moral*. New York: Oxford University Press; 1999. p. 134.

Correspondence to: *Dr. Olli S. Miettinen, Department of Epidemiology and Biostatistics, Faculty of Medicine, McGill University, 1020 Pine Ave. West, Montreal QC H3A 1A2; fax 514 398-4503; osm@epid.lan.mcgill.ca*

Articles to date in this series

Miettinen OS. The modern scientific physician: 1. Can practice be science? *CMAJ* 2001;165(4):441-2.
Miettinen OS. The modern scientific physician: 2. Medical science versus scientific medicine. *CMAJ* 2001; 165(5):591-2.
Miettinen OS. The modern scientific physician: 3. Scientific diagnosis. *CMAJ* 2001;165(6):781-2.