Algorithms and surrogate markers in translational research

Paul van Helden

Most clinicians are familiar with the concept of diagnostic markers. A physician takes a sample—urine, blood, saliva or tissue—and gives it to an analytical laboratory that returns either a positive or negative result that might or might not lead to a diagnosis of disease. In fact, however, the real result is not qualitative but quantitative: the concentration of cholesterol, an inflammation marker or blood sugar, for example. It is turned into a qualitative measure by comparing it with an empirically decided ‘normal’ value; that is, it is either higher or lower than the ‘norm’. As scientists, we ought to know that such analyses are subject to errors and outliers, and that the ‘norm’ itself can mean little when applied to certain individuals. In the clinical assessment of a patient, then, the clinician combines the symptoms, the laboratory results and his or her own experience or intuition to make a diagnosis.

The basis of the diagnosis is effectively an algorithm, even if it does not use binary logic: symptom 1 indicates several possible options, which are further reduced by analytical test 1 and symptom 2 and so on, until the algorithm settles down to a diagnostic decision. Unfortunately, there are many symptoms in clinical medicine that are common to different ailments and there is great heterogeneity among patients, which complicates decision-making. Things are made even harder when the patient cannot communicate with the clinician, as is the case with infants or animals. To reach a definitive diagnosis is therefore often difficult, even when aided by laboratory analysis. Moreover, some established laboratory assays have turned out to be questionable predictors of disease, such as the prostate-specific antigen (PSA) test for which a high value does not necessarily indicate prostate cancer.

Nevertheless, we develop instruments and assays that offer ever-greater precision and accuracy. Yet, when ‘hard data’ is clinically interpreted against a noisy background, are greater precision and accuracy useful? Part of the answer involves the method we use to calculate the acceptable level of accuracy for any test. For example, by using a Bayesian approach, even if a given assay is 99% sensitive and specific, the chance that a positive test predicts the condition is just 50% if only 1% of the population is affected.

However, several tests, even with relatively low sensitivity and specificity, can combine to provide a high confidence level for decision-making. If we know that 50% of individuals with a certain quantitative value have a given disease, and combine this with positive results from other tests that also have 50% certainty, we can reach a far higher confidence level. Each test alone has little meaning, but combined they become highly informative. This implies that we need to refine laboratory assays to the highest possible accuracy—sensitivity and specificity—for use in diagnostic algorithms.

The use of such algorithms could go beyond their application to diagnosis. Most tests measure some parameter that is directly related to the condition, such as PSA in prostate cancer or cholesterol in cardiovascular disease. However, we might use other markers that are not apparently related to the immediate condition—so-called ‘surrogates’. Such markers could be used for diagnosis, but they have other potentially valuable uses: to monitor the outcome of drug trials or vaccine trials, or to monitor the patient’s response to treatment, for example. The key is to decide which combination of surrogate markers should be used in each situation, and what weight each should be given in each case. This requires a large

sample base and extensive testing to ensure that the best quantitative measures are used to set standard baselines.

Surrogates could be particularly useful for clinical drug trials with lengthy patient follow-up, particularly when the efficacy measure includes the rate of relapse or adverse effects over a period of years after therapy. It is not possible to predict the clinical outcome of a new treatment or to predict whether a patient will relapse after therapy, so a surrogate marker that provides an early indication of efficacy would aid patients and should decrease the time needed to conduct clinical trials. Moreover, surrogate markers could be used to identify the patients most suitable for the trial in the first place, or patients who will respond poorly to standard therapy and who might benefit from an alternative. They might also help to identify patients at high risk of relapse after treatment.

New tests for surrogate markers should therefore help clinicians to improve the efficiency of drug trials, thereby reducing cost. Even if a test is not highly sensitive and specific, we should not discount its value; it might be useful for developing algorithms that use a growing tool kit of surrogate markers. In fact, any additional test has a multiplicative effect on its usefulness. Surrogate markers and refined diagnostic algorithms have huge potential to improve drug testing and to help clinicians make rapid and accurate diagnoses.

Paul van Helden is a professor and Head of the Department of Biomedical Science at Stellenbosch University, Faculty of Health Science, Tygerberg, Cape Town, South Africa. E-mail: pvh@sun.ac.za

Published online 18 November 2011

EMBO reports (2011) 12, 1206.
doi:10.1038/embor.2011.215