Please, Be Sensitive, Specific, Discriminative, and Predictive ...

Stefano Ricci
UO Neurologia, USL Umbria 1, Città di Castello, Italy

This time, the methodological issues we are going to discuss here have to do with diagnosis, that is the process we stroke clinicians are endlessly involved in as we practice our profession. Anything we ourselves do or ask for (like taking a patient’s history or examining him or her, evaluating blood test results, looking at CT or MR images, performing Doppler ultrasound, etc.) needs to be done accurately if we are to ‘dispel the fog’ which conceals the correct diagnosis for a particular patient. But the word ‘accurate’ is not synonymous with ‘100% exact’, simply because this level of precision is rarely, if ever, possible in real life. We therefore need to know what we risk if one of our diagnostic tools (from the simplest to the most technologically sophisticated) leads us to a wrong conclusion. The paper by Eusebi [1] helps us understand how to measure this risk of being wrong, using epidemiological data like sensitivity, specificity, and predictive values. But, as perhaps the most important statement in the paper suggests, ‘how much discriminative or predictive power we need depends on the clinical diagnostic pathway and on the misclassification costs’. To me, this statement means that, once again, it is up to us clinicians to guide the use of statistical and epidemiological tools, with the goal of addressing the problems afflicting every individual and very special patient we treat. We cannot simply say to the test: ‘please, be sensitive, specific, discriminative, and predictive’; we have to choose when it would be better to live with some deficiency in (say) a positive predictive value in order to have a very high negative predictive value or vice versa. For instance, to miss atrial fibrillation in stroke patients is much worse (in terms of the available prevention possibilities) than to miss an asymptomatic carotid stenosis in a patient without a clear history of transient ischemic attack or stroke. However, we could certainly use our diagnostic resources better if we knew how to measure their potential with the various instruments Eusebi clearly describes and demonstrates in his paper.

Needless to say, as always in medicine, it may happen that we find some data about accuracy results for a given diagnostic procedure in one paper and slightly different data in another paper. Although we know that disease prevalence can influence some of these results (namely, predictive values), we may wonder how to deal with these differences. Would it be possible to ‘sum up’ the results in a systematic review as we do with therapeutic measures? This will be the central topic of the next paper in the methodological section.

Reference