Comparative Analysis of Different Approaches to Report Diagnostic Accuracy

The accuracy of a diagnostic test is traditionally reported in terms of sensitivity and specificity and predictive values in scientific articles, systematic reviews, and test inserts. Physicians tend, however, to confuse sensitivity and specificity, which are independent of the prevalence, with positive and negative predictive values, which are dependent on the prevalence.1

An alternative way of expressing diagnostic accuracy is the likelihood ratio (LR).2 Likelihood ratios offer the advantage that they are not dependent on the prevalence of the disease, in contrast to positive and negative predictive values, and can be used for nondichotomous test results.3,4 Steurer and coworkers1 found that reporting diagnostic accuracy as LRs expressed in nontechnical language (Figure 1) offers the advantage that they are not dependent on the prevalence and results of diagnostic accuracy. Depiction of the posttest probability as a function of the prevalence is a graphic representation that can be understood more easily by physicians.5

Most practicing physicians do not use the Bayesian approach to calculate posttest probability because they consider it impractical and difficult to calculate.7 The use of a graphic representation of the posttest probability as a function of the pretest probability could overcome these problems. This approach has, however, not been studied. While such approach is ill suited for printed laboratory reports, the current advances in information technology allow laboratories to provide this graphic representation “one mouse click away” to clinicians when results are checked electronically.

Methods. We developed a 3-item questionnaire. Each question asked the respondent to estimate the posttest probability and included a pretest probability in the form of a prevalence and results of diagnostic accuracy. Depending on the question, the diagnostic accuracy was presented as sensitivity and specificity, LR in nontechnical terms, or graphically. The questionnaire, in English translation, is shown in Figure 1.

The questionnaire was distributed to 117 general practitioners and 55 specialists in internal medicine, including trainees, at 2 separate conferences on continuing medical education. The program did not include any specific information regarding the performance of diagnostic tests. The chairman asked participants to complete the questionnaire during the small breaks between the speakers. The questionnaires were collected immediately after the last speaker of the session.

Of the 172 eligible clinicians, 123 filled in the questionnaire (response rate, 72%). The clinicians answering the questions had on average more than 25 years of experience as a physician and had on average more than 40 patient contact hours per week. The correct answer to question 1 was “approximately 20%” (19.6% exactly). The second question described the same test as the first question, but the diagnostic accuracy information was described as an LR in nontechnical language (Figure 1). Since the LR equals sensitivity/(1 – specificity), the LR for the first test was 9.5. The correct answer for question 2 was also “approximately 20%” (20.4% exactly).

The third question provided a graphic representation of the posttest probability as a function of the pretest probability. The test depicted was relatively similar to the tests in the 2 previous questions, with an LR of 12.4 for a positive result. Given the higher prevalence of 20%, the correct answer was “approximately 80%” (75.5%...
negative result

Figure 2. A graphic depiction of pretest vs posttest probability of the disease. Using this type of graph, one can calculate the posttest probability of the disease for any given pretest probability for a specific laboratory test. For example, if the pretest probability is 10% (the patient has a 10% chance of having the target disorder) and the depicted test result is positive, then the posttest probability of the disease is 60% (the patient has a 60% chance of having the target disorder if the test result is positive).

exactly). The test depicted in question 3 is an IgA anti-tissue transglutaminase assay used in our laboratory (diagnostic accuracy in routine clinical practice).

The proportion of physicians who gave a correct answer was calculated for each of the 3 questions. Physicians who indicated that they did not know the answer to a particular question were not included for this calculation. A Fisher exact test was used to evaluate significant differences in proportions. Calculations were performed using GraphPad QuickCalcs (http://www.graphpad.com) from GraphPad Software (GraphPad Software Inc, La Jolla, California).

Results. Most clinicians grossly overestimated the probability of disease when the diagnostic accuracy information was given as sensitivity and specificity (question 1). Most clinicians (81%) indicated that the probability of disease was approximately 90%. The response by younger (≤40 years old) and older physicians (≥41 years old) was comparable (7% vs 8% correct answers, respectively).

When the diagnostic accuracy information of the same test was provided as an LR in nontechnical terms (question 2), the number of correct answers increased significantly from 8% to 34% (P < .001; Figure 1). Of the 93 clinicians who answered questions 1 and 2, 43 (46%) gave the same estimate twice, 41 (44%) gave a lower estimate for question 2 than for question 1, and only 9 (10%) gave a higher estimate for question 2 than for question 1.

With a graphic approach (question 3), 73% of the responding clinicians gave the correct answer (P < .001 vs questions 1 and 2). The number of nonresponders for question 3 (38%) was higher than for question 1 (15%) (P = .002) and question 2 (22%) (P = .03).

Comment. Our study demonstrates that a graphic approach is the most efficient way to convey diagnostic accuracy information to clinicians. The lack of familiarity with a graphic approach, indicated by the higher proportion of persons who could not answer question 3 compared with question 1, is hardly surprising, since previous studies found that most clinicians also lack familiarity with LRs. Because the graphic approach does not require calculations to estimate the posttest probability, in contrast to the use of sensitivity and specificity or LRs, we believe that the graphic approach is much easier to explain to clinicians (Figure 2). Since most physicians in industrialized countries consult their laboratory results electronically, a graphic representation could be provided “one mouse click away” to clinicians.

Pieter Vermeersch, MD, PhD
Xavier Bossuyt, MD, PhD

Author Affiliations: Laboratory Medicine, University Hospitals Leuven, Leuven, Belgium.

Correspondence: Dr Vermeersch, Laboratory Medicine, University Hospitals Leuven, Herestraat 49, B-3000 Leuven, Belgium (pieter.vermeersch@uz.kuleuven.be).

Author Contributions: Study concept and design: Vermeersch and Bossuyt. Acquisition of data: Vermeersch. Analysis and interpretation of data: Vermeersch. Drafting of the manuscript: Vermeersch. Critical revision of the manuscript for important intellectual content: Vermeersch and Bossuyt. Statistical analysis: Vermeersch. Study supervision: Bossuyt.

Financial Disclosure: None reported.


COMMENTS AND OPINIONS

Progression of Coronary Artery Calcification: Not Down-and-Out

In the December 14, 2009, issue, McCullough et al1 presented a systematic review of randomized trials that used coronary artery calcification (CAC) progression as the primary end point. Despite the noninclusion of 2 eligible randomized placebo-controlled trials,2,3 we agree with their assertion that CAC progression has not been consistently reduced in any therapeutic randomized trial to date. The field of CAC progression has been hampered by the issue of variability between scans and a lack of con-