Objective assessment of diagnostic tests validity: a short review for clinicians and other mortals. Part II

Nermin N. Salkić

University Clinical Center Tuzla
Department of Gastroenterology
Bosnia and Herzegovina

Corresponding author:
Nermin Salkić
Miroslava Krleže 15/26
75000 Tuzla,
Bosnia and Herzegovina
snermin@gmail.com

Received: 30 November 2008
Accepted: 23 December 2008

The whole point of a diagnostic test is to use it to make a diagnosis, thus the obvious need is to know how accurately a particular diagnostic test detects patients with or without a disease. In order to know it, a clinician or a researcher should have a basic understanding of the principles of objective appraisal of diagnostic test. In the second part of this short review, the author presents the most common biostatistical methodology for assessment of a validity of diagnostic tests. Definitions and interpretations of accuracy and likelihood ratio are also provided together with methods of their calculation.

Key words: Accuracy, Likelihood ratio.

Introduction

In the previous part we discussed sensitivity, specificity and positive and negative predictive values. We are continuing with an explanation of accuracy and likelihood ratio. In order to make calculation easier, let us remind ourselves of our imaginary clinical research.

As previously stated, liver biopsy is currently considered to be the gold standard in the assessment of the presence and degree of liver fibrosis in various liver diseases, such as viral hepatitis etc (1). However, it is associated with the possibility of severe complications and serious discomfort for the patient (2). Therefore, our hypothetical investigators decided to evaluate a non-invasive marker of liver fibrosis comparing it against the gold standard (liver biopsy).

Investigators recruited 189 patients. After performing a liver biopsy, 43 of them had liver fibrosis, while 146 did not. On the other hand, after performing a non-invasive test for liver fibrosis, 61 patients were positive for the presence of liver fibrosis, while 128 of them were negative. Now, let us make a 2-by-2 table out of this data (Table 1).

Accuracy

If we want to know the overall accuracy of a test we will need to calculate what proportion of all tests have given the correct result (true positives and true negatives as a pro-

Table 1 Two by two table showing the results of validation study of non-invasive liver fibrosis test against gold standard

	Liver biopsy positive	Liver biopsy negative	Total
Test positive	43 (TP)	18 (FP)	61
Test negative	0 (FN)	128 (TN)	128
Total	43	146	189

TP-true positive; TN-true negative; FP-false positive; FN-false negative

portion of all results) ? (3). For this particular task we will use the formula

Accuracy = (TP+TN)/(TP+TN+FP+FN), or in our case

Accuracy=(43+128)/(43+128+18+0)= 171/189=0,905=90,5%

This means that our test correctly classifies 90.5% of patients (true positives and negatives).

Likelihood ratios

Although the sensitivity and specificity of a test are virtually constant whatever the prevalence of the condition, the positive and negative predictive values depend crucially on prevalence (prevalence in our study=43/189=0,2275=22,75%). When we change the prevalence PPV and NPV change also. The lower the prevalence the more sure we can be that a negative test result indicates the absence of a condition and the less sure we can be that a positive result really indicate the presence of a condition (4).

In order to avoid the impact of prevalence and to express the usefulness of a diagnostic test, likelihood ratios can be calculated. Likelihood ratio (LR) expresses the magnitude by which the probability of a diagnosis in a given patient is modified by the result of a test (5). LR for a test result is the ratio between the chance of observing that result in a patient with the disease in question and the chance of that result in subjects without the disease. The likelihood ratio of a positive test (LR+) answers the question "How much more likely is a positive test to be found in a person with the condition than in a person without it?" (3). It is calculated by using the formula

(LR+)=Sensitivity/(100-Specificity), or in our case

(LR+)=100 /(100-88)=100/12=8.33.

On the other hand, the likelihood ratio of a negative test (LR-) answers the question "How much more likely is a negative test to be found in a person without the condition than in a person with it?" (3) It is calculated by using the formula

(LR-)=(100-Sensitivity)/Specificity, or in case of our hypothetical study

(LR-)=(100-100)/88=0/88=0.

How to use likelihood ratios

Now that we know how to calculate LR, how do we use them? Well, we need the help of another formula:

Post-test odds= Pre-test odds of disease x likelihood ratio

We know that in our imaginary research, the prevalence (or pre-test probability) of liver fibrosis is 22.7%. But, imagine that you are working in a clinical setting where the prevalence of liver fibrosis is higher, say 33%. How useful is the non-invasive liver fibrosis test in this case? First we need to calculate pre-test odds from probability, which is a simple task:

Odds=probability/1-probability=0,33/1-0,33=0,33/0,67=1/2

Now, we calculate the post-test odds for liver fibrosis:

Post-test odds = Pre-test odds of disease x likelihood ratio = 1/2 x 8,33 = 8,33/2 = 4,16/1

Post test odds of liver fibrosis are therefore "4.16 in favor to 1 against". We have the post-test odds, so all we have to do is to convert back to post-test probability using the following formula:

Probability = odds in favor/odds in favor + odds against=
=4,16/4,16+1=4,16/5,16=0,806=80,6%

So, after a patient tested positive with our non-invasive liver fibrosis test in this particular clinical setting, we can conclude that he has an 80.6% chance of actually having liver fibrosis. The rationale for calculating the (LR-) is the same. In order to help in the process of calculating post-test probabilities, we are providing here a nomogram from an original paper by Fagan, for working out post-test probabilities when the pretest probability (prevalence) and likelihood ratio for the test are known (Figure 1) (6).

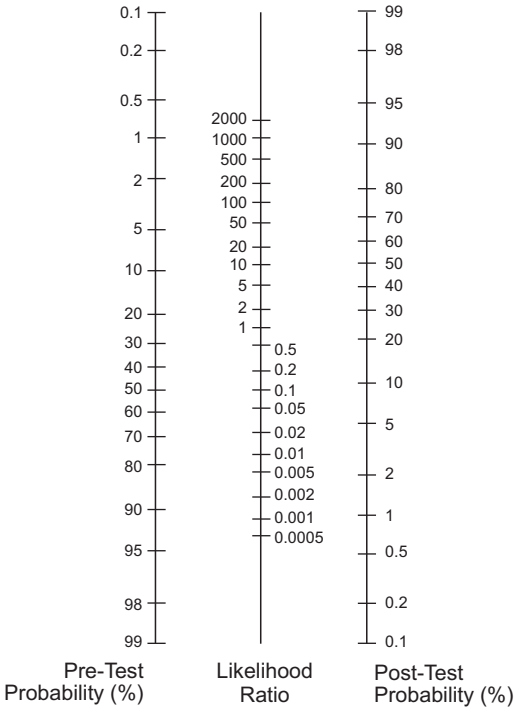

Figure 1 Likelihood ratio nomogram for easy calculation of pre-test and post-test probabilities based on likelihood ratio values

When we have likelihood ratios, the very high or very low prevalence of disease (or pre-test probabilities) are less likely to influence the post-test probability of disease. Although one may question the objectivity of selecting the pre-test probabilities, combining the objectivity of likelihood ratios with subjective pretest probabilities is indeed consistent with the principles of evidence based medicine(7). Although the concept of LR is somewhat hard to grasp, it is of tremendous help in selecting the appropriate diagnostic test based on published results regarding their validity, but in our own clinical setting. They can be used to combine several diagnostic tests; they can be calculated for several threshold values of each particular test (7). An LR greater than 1 gives a post-test probability which is higher than the pre-test probability. An LR less than 1 produces a post-test probability which is lower than the pre-test probability. How much LR actually changes disease likelihood is presented in Table 2.

Table 2 Impact of likelihood ratio values on likelihood of a disease

High LR	Low LR	Impact on Likelihood
>10	<0,1	Large
5-10	0,1-0,2	Moderate
2-5	0,2-0,5	Small
<2	>0,5	Tiny
1	1	No Change

The likelihood ratio thus has enormous practical value, and it is becoming the preferred way of expressing and comparing the usefulness of different tests (3). Moreover, the likelihood concept is applicable in many other situations. Since the likelihood ratio is the ratio of the maximum probability of a result under two different hypotheses, often a null hypothesis and an alternative hypothesis, it can be used as a statistical test (likelihood ratio test) for making a decision

between two hypotheses based on the value of this ratio. It is also possible to use the likelihood concept for calculation of confidence intervals, comparison of two groups, regression models etc, the details of which are well beyond the scope of this article.

We hope that this short review provides the basic information and explanation necessary for a busy clinician to become acquainted with the methodology of diagnostic test assessment. We will continue our series on biostatistics in our next issue with discussion on a different subject.

References

1. Lok AS, McMahon BJ. Chronic hepatitis B. Hepatology. 2001;34:1225-41.
2. Cadranel JF, Rufat P, Degos F. Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AFEF). Hepatology. 2000;32:477-81.
3. Greenhalgh T. How to read a paper. Papers that report diagnostic or screening tests. BMJ. 1997;315:540-3.
4. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. BMJ. 1994;309:102.
5. Halkin A, Reichman J, Schwaber M, Paltiel O, Brezis M. Likelihood ratios: getting diagnostic testing into perspective. QJM. 1998;91(4):247-58.
6. Fagan TJ. Letter: Nomogram for Bayes theorem. N Engl J Med. 1975;293:257.
7. Okada T, Rao G. Using the likelihood ratio. J Fam Pract. 2005;54:127-8.