Two issues were simplified [letter]

ter Riet, G.; Kessels, A.G.H.; Bachmann, L.M.

Published in:
BMJ : British medical journal

Citation for published version (APA):
Systematic reviews of evaluations of diagnostic and screening tests

Odds ratio is not independent of prevalence

Two issues were simplified

Odds ratio is not independent of prevalence

EDITOR—Deeks, in the third of four articles on evaluations of diagnostic and screening tests, promoted the odds ratio as often being constant regardless of the diagnostic threshold.† We agree with Deeks's statement that the choice of threshold varies according to the prevalence of the disease. But the statement that the odds ratio is generally constant regardless of the diagnostic threshold can be misleading. The value of an odds ratio, like that of other measures of test performance—for example, sensitivity, specificity, and likelihood ratios—depends on prevalence.‡ For example, a test with a diagnostic odds ratio of 10.00 is considered to be a very good test by current standards. It is easy to verify that this is generally true only in populations at high risk. A diagnostic odds ratio of 10.00 in a low risk population may represent a very weak association between the experimental test and the gold standard test. This is so because the observable range of values for an odds ratio increases as the prevalence of the disease decreases (moves away from 1/2).

Nicole Jill-Marie Blackman, senior biostatistician.
GlaxoSmithKline, 1250 South Collegeville Road, PO Box 5089, Collegeville, PA 19426-0989, USA
nicole_blackman-1@gsk.com

EDITOR—We would like to draw attention to two points, which Deeks in his review article simplified.

Firstly, consider an example to illustrate the futility of what might be called the "reflex to fill the fourfold table" in research into diagnostic accuracy. Consider a study on an experimental test that claims to give clinicians more certainty in situations where they have only a few indications that disease may be present. But let us assume that the indications are not strong enough to justify the performance of truly invasive tests. Without the new experimental test these patients would be sent home. The value of the new test lies in its ability to identify those patients who have the disease and would benefit from treatment. In this scenario, the analysis of only those patients who test positively on the experimental test (two cells filled of the fourfold table) suffices to learn about its usefulness.

Secondly, Deeks ends his explanation of the application of the likelihood ratio by saying that knowledge of other characteristics of a particular patient that either increase or decrease their prior probability of endometrial cancer can be incorporated into the calculation by adjusting the pretest probability accordingly. This, however, assumes constancy of likelihood ratios, which should not be assumed because it is usually incorrect. In practice, the knowledge of other patient characteristics will have an influence on the magnitude of the likelihood ratios of following tests. This is so because when a chain of diagnostic tests (history taking, physical exam, lab tests, or imaging) is performed on a patient, certain results from his or her clinical history make the likelihood to find certain lab results more (or less) likely. This in turn influences the chances of finding certain imaging results. In other words, the results of the component tests are not mutually independent. For example, on average, women with a positive test on ultrasound (thickened endometrium) are more likely to test positively on hysteroscopy, in which the endometrial thickness is also assessed, albeit in a different manner.

The theoretical solution to this problem is the calculation of likelihood ratios that are conditional on the results of the preceding tests in the diagnostic test chain. In practice, this is usually not feasible owing to lack of data, and most investigators use logistic regression models to account for all these dependencies. These models, however, yield diagnostic odds ratios, not likelihood ratios. It is partly this complexity that hampers the application of simple diagnostic accuracy studies to clinical practice.
In figure 2 the numerators in the second column of the right hand panel represent the number of false positives, not the true negatives.

**Gerben ter Riet**, *clinical epidemiologist.*
Department of General Practice, Academic Medical Centre, University of Amsterdam, 1105 AZ Amsterdam, Netherlands g.terriet@epid.unimaas.nl

**Alphons G H Kessels**, *medical statistician.*
Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Hospital, 6202 AZ Maastricht, Netherlands

**Lucas M Bachmann**, *research fellow.*
Horten Centre, University of Zurich, CH-8091 Zurich, Switzerland


© BMJ 2001

**Related Article**

*Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests*
Jonathan J Deeks
BMJ 2001 323: 157-162. [Extract] [Full Text]