

## General practice

### Statistics Notes: Diagnostic tests 2: predictive values

**Douglas G Altman**, *head Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE*,<sup>a</sup> **J Martin Bland**, *reader in medical statistics*<sup>a</sup>

<sup>a</sup> Medical Statistics Laboratory, Imperial Cancer Research Fund, London WC2A 3PX

The whole point of a diagnostic test is to use it to make a diagnosis, so we need to know the probability that the test will give the correct diagnosis. The sensitivity and specificity<sup>1</sup> do not give us this information. Instead we must approach the data from the direction of the test results, using predictive values.

Positive predictive value is the proportion of patients with positive test results who are correctly diagnosed.

Negative predictive value is the proportion of patients with negative test results who are correctly diagnosed.

Using the same data as in the previous note,<sup>1</sup> we know that 231 of 263 patients with abnormal liver scans had abnormal pathology, giving the proportion of correct diagnoses as  $231/263 = 0.88$ . Similarly, among the 81 patients with normal liver scans the proportion of correct diagnoses was  $54/81 = 0.59$ . These proportions are of only limited validity, however. The predictive values of a test in clinical practice depend critically on the prevalence of the abnormality in the patients being tested; this may well differ from the prevalence in a published study assessing the usefulness of the test.

This is the fourth in a series of occasional notes on medical statistics.

In the liver scan study the prevalence of abnormality was 0.75. If the same test was used in a different clinical setting where the prevalence of abnormality was 0.25 we would have a positive predictive value of 0.45 and a negative predictive value of 0.95. The rarer the abnormality the more sure we can be that a negative test indicates no abnormality, and the less sure that a positive result really indicates an abnormality. Predictive values observed in one study do not apply universally.

The positive and negative predictive values (PPV and NPV) can be calculated for any prevalence as follows:

$$\text{PPV} = \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1 - \text{specificity}) \times (1 - \text{prevalence})}$$

$$\text{NPV} = \frac{\text{specificity} \times (1 - \text{prevalence})}{(1 - \text{sensitivity}) \times \text{prevalence} + \text{specificity} \times (1 - \text{prevalence})}$$

If the prevalence of the disease is very low, the positive predictive value will not be close to 1 even if both the sensitivity and specificity are high. Thus in screening the general population it is inevitable that many people with positive test results will be false positives.

The prevalence can be interpreted as the probability before the test is carried out that the subject has the disease, known as the prior probability of disease. The positive and negative predictive values are the revised estimates of the same probability for those subjects who are positive and negative on the test, and are known as posterior probabilities. The difference between the prior and posterior probabilities is one way of assessing the usefulness of the test.

For any test result we can compare the probability of getting that result if the patient truly had the condition of interest with the corresponding probability if he or she were healthy. The ratio of these probabilities is called the likelihood ratio, calculated as sensitivity/ (1 - specificity).

The likelihood ratio indicates the value of the test for increasing certainty about a positive diagnosis. For the liver scan data the prevalence of abnormal pathology was 0.75, so the pre-test odds of disease were  $0.75/(1 - 0.75) = 3.0$ . The sensitivity was 0.895 and the specificity was 0.628. The post-test odds of disease given a positive test is  $0.878/(1 - 0.878) = 7.22$ , and the likelihood ratio is  $0.895/(1 - 0.628) = 2.41$ . The posttest odds of having the disease is the pre-test odds multiplied by the likelihood ratio.

A high likelihood ratio may show that the test is useful, but it does not necessarily follow that a positive test is a good indicator of the presence of disease.

## Related Article

### Diagnostic accuracy of clinical examination for detection of non-cephalic presentation in late pregnancy: cross sectional analytic study

Natasha Nassar, Christine L Roberts, Carolyn A Cameron, and Emily C Olive

*BMJ* 2006 333: 578-580. [Abstract] [Full Text]

### This article has been cited by other articles:

(Search Google Scholar for Other Citing Articles)

- Gorricho, J., Erviti, J., J. Henry, M., Pasco, J. A., Sanders, K. M., Nicholson, G. C., Kotowicz, M. A. (2007). Fracture Risk Score. *Radiology* 245: 298-298 [Full text]
- Madlensky, L., Daftary, D., Burnett, T., Harmon, P., Jenkins, M., Maskiell, J., Nigon, S., Phillips, K., Templeton, A., Limburg, P. J., Haile, R. W., Potter, J. D., Gallinger, S., Baron, J. A. (2007). Accuracy of Colorectal Polyp Self-Reports: Findings from the Colon Cancer Family Registry. *Cancer Epidemiol. Biomarkers Prev.* 16: 1898-1901 [Abstract] [Full text]
- Affolabi, D., Odoun, M., Martin, A., Palomino, J. C., Anagonou, S., Portaels, F. (2007). Evaluation of Direct Detection of Mycobacterium tuberculosis Rifampin Resistance by a Nitrate Reductase Assay Applied to Sputum Samples in Cotonou, Benin. *J. Clin. Microbiol.* 45: 2123-2125 [Abstract] [Full text]
- Mausbach, B. T., Harvey, P. D., Goldman, S. R., Jeste, D. V., Patterson, T. L. (2007). Development of a Brief Scale of Everyday Functioning in Persons with Serious Mental Illness. *Schizophr Bull* 0: sbm014v1-1 [Abstract] [Full text]
- Cucchetti, A, Vivarelli, M, Heaton, N D, Phillips, S, Piscaglia, F, Bolondi, L, La Barba, G, Foxton, M R, Rela, M, O'Grady, J, Pinna, A D (2007). Artificial neural network is superior to MELD in predicting mortality of patients with end-stage liver disease. *Gut* 56: 253-258 [Abstract] [Full text]
- Lex, C., Ferreira, F., Zacharasiewicz, A., Nicholson, A. G., Haslam, P. L., Wilson, N. M., Hansel, T. T., Payne, D. N. R., Bush, A. (2006). Airway Eosinophilia in Children with Severe Asthma: Predictive Values of Noninvasive Tests. *Am. J. Respir. Crit. Care Med.* 174: 1286-1291 [Abstract] [Full text]
- Lozano-Calderon, S., Blazar, P., Zurakowski, D., Lee, S.-G., Ring, D. (2006). Diagnosis of Scaphoid Fracture Displacement with Radiography and Computed Tomography. *J. Bone Joint Surg. Am.* 88: 2695-2703 [Abstract] [Full text]
- Nassar, N., Roberts, C. L., Cameron, C. A., Olive, E. C (2006). Diagnostic accuracy of clinical examination for detection of non-cephalic presentation in late pregnancy: cross sectional analytic study. *BMJ* 333: 578-580 [Abstract] [Full text]
- Hoffmann, S, Hartung, T (2006). Toward an evidence-based toxicology. *Hum Exp Toxicol* 25: 497-513 [Abstract]
- Luo, H-D, Gazzard, G, Liang, Y, Shankar, A, Tan, D T., Saw, S-M (2006). Defining myopia using refractive error and uncorrected logMAR visual acuity >0.3 from 1334 Singapore school children ages 7-9 years.. *Br. J. Ophthalmol.* 90: 362-366 [Abstract] [Full text]
- Rajappan, K., Murphy, E., Amber, V., Meakin, F., Muller, B., Fox, K.F., Baker, C.S.R. (2005). Usage of troponin in the real world: a lesson for the introduction of biochemical assays. *QJM* 98: 337-342 [Abstract] [Full text]
- Mekontso-Dessap, A., Honore, S., Kirsch, M., Houel, R., Loisanee, D., Brun-Buisson, C. (2004). Usefulness of Routine Epicardial Pacing Wire Culture for Early Prediction of Poststernotomy Mediastinitis. *J. Clin. Microbiol.* 42: 5245-5248 [Abstract] [Full text]
- Deeks, J. J, Altman, D. G (2004). Diagnostic tests 4: likelihood ratios. *BMJ* 329: 168-169 [Full text]
- Hippisley-Cox, J., Hammersley, V., Pringle, M., Coupland, C., Crown, N., Wright, L. (2004). Methodology for Assessing the Usefulness of General Practice Data for Research in One Research Network. *Health Informatics Journal* 10: 91-109 [Abstract]
- Bunn, H.J., Woltmann, G., Grigg, J. (2002). Applicability of laser scanning cytometry to study paediatric alveolar macrophages. *Eur Respir J* 20: 1437-1443 [Abstract] [Full text]
- Equi, A C, Pike, S E, Davies, J, Bush, A (2001). Use of cough swabs in a cystic fibrosis clinic. *Arch. Dis. Child.* 85: 438-439 [Abstract] [Full text]
- Chinn, S. (2001). Statistics for the European Respiratory Journal. *Eur Respir J* 18: 393-401 [Abstract] [Full text]
- Hassey, A., Gerrett, D., Wilson, A. (2001). A survey of validity and utility of electronic patient records in a general practice. *BMJ* 322: 1401-1405 [Abstract] [Full text]
- McKeith, I. G., Ballard, C. G., Perry, R. H., Ince, P. G., O'Brien, J. T., Neill, D., Lowery, K., Jaros, E., Barber, R., Thompson, P., Swann, A., Fairbairn, A. F., Perry, E. K. (2000). Prospective validation of Consensus criteria for the diagnosis of dementia with Lewy bodies. *Neurology* 54: 1050-1058 [Abstract] [Full text]
- Foster, P. J, Devereux, J. G, Alsibirk, P. H., Lee, P. S., Uranchimeg, D., Machin, D., Johnson, G. J, Baasanhu, J. (2000). Detection of gonioscopically occludable angles and primary angle closure glaucoma by estimation of limbal chamber depth in Asians: modified grading scheme. *Br. J. Ophthalmol.* 84: 186-192 [Abstract] [Full text]
- Hussaini, S H, Sheridan, M B, Davies, M (1999). The predictive value of transabdominal ultrasonography in the diagnosis of biliary tract complications after orthotopic liver transplantation. *Gut* 45: 900-903 [Abstract] [Full text]
- Tsuang, D., Larson, E. B., Bowen, J., McCormick, W., Teri, L., Nochlin, D., Leverenz, J. B., Peskind, E. R., Lim, A., Raskind, M.

- A., Thompson, M. L., Mirra, S. S., Gearing, M., Schellenberg, G. D., Kukull, W. (1999). The Utility of Apolipoprotein E Genotyping in the Diagnosis of Alzheimer Disease in a Community-Based Case Series. *Arch Neurol* 56: 1489-1495 [Abstract] [Full text]
- Maurer, M., Shambal, S., Berg, D., Woydt, M., Hofmann, E., Georgiadis, D., Lindner, A., Becker, G. (1998). Differentiation Between Intracerebral Hemorrhage and Ischemic Stroke by Transcranial Color-Coded Duplex-Sonography. *Stroke* 29: 2563-2567 [Abstract] [Full text]
  - McConville, J P, Craig, J J, Collinge, J., Rossor, M. N, Thomas, D., Frosh, A., Tolley, N., Otto, M., Zerr, I., Poser, S., Wiltfang, J., Schütz, E., Pfahlberg, A., Gefeller, O. (1998). Diagnosis of Creutzfeldt-Jakob disease by measurement of S100 protein in serum. *BMJ* 317: 472-472 [Full text]
  - Maffulli, N. (1998). The Clinical Diagnosis of Subcutaneous Tear of the AchillesTendon: A Prospective Study in 174 Patients. *Am J Sports Med* 26: 266-270 [Abstract] [Full text]
  - Altman, D G, Bland, J M (1994). Statistics Notes: Diagnostic tests 3: receiver operating characteristic plots. *BMJ* 309: 188-188 [Full text]