

Models In Epidemiology And Biostatistics  
Gordon Hilton Fick

Sensitivity, Specificity and Predictive Value  
[adapted from Altman and Bland – BMJ.com]

The simplest diagnostic test is one where the results of an investigation, such as an x ray examination or biopsy, are used to classify patients into two groups according to the presence or absence of a symptom or sign. For example, the table shows the relation between the results of a test, a liver scan, and the correct diagnosis based on either necropsy, biopsy, or surgical inspection. How good is the liver scan at diagnosis of abnormal pathology?

Relation between results of liver scan and correct diagnosis

Liver scan	Pathology		Total
	Abnormal (+)	Normal (-)	
Abnormal (+)	231	32	263
Normal (-)	27	54	81
Total	258	86	344

One approach is to calculate the proportions of patients with normal and abnormal liver scans who are correctly "diagnosed" by the scan. The terms positive and negative are used to refer to the presence or absence of the condition of interest, here abnormal pathology. Thus there are 258 true positives and 86 true negatives. The proportions of these two groups that were correctly diagnosed by the scan were  $231/258=0.90$  and  $54/86=0.63$  respectively. These two proportions are estimates of probabilities.

The sensitivity [Sn] of a test is the probability that the test is positive given a patient has the condition.

$$Sn = P(T+ | D+)$$

The specificity [Sp] of a test is the probability that the test is negative given a patient does not have the condition.

$$Sp = P(T- | D-)$$

We can thus say that, based on the sample studied, we would estimate that 90% of patients with abnormal pathology would have abnormal (positive) liver scans, while 63% of those with normal pathology would have normal (negative) liver scans.

$$\hat{Sn} = 0.8953 \text{ and } \hat{Sp} = 0.6279$$

The sensitivity and specificity are probabilities, so confidence intervals can be calculated for them using standard methods for proportions.

Using Stata: ( cii is confidence interval immediate )

```
. cii 258 231
```

Variable	Obs	Mean	Std. Err.	-- Binomial Exact -- [95% Conf. Interval]	
	258	.8953488	.0190572	.8513977	.9298934

```
. cii 86 54
```

Variable	Obs	Mean	Std. Err.	-- Binomial Exact -- [95% Conf. Interval]	
	86	.627907	.0521224	.5169596	.7297749

Sensitivity and specificity are one approach to quantifying the diagnostic ability of the test. In clinical practice, however, the test result is all that is known, so we want to know how good the test is at predicting abnormality. In other words, what is the probability that a patient with abnormal test results is truly abnormal?

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 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 probability that a patient with abnormal test results is truly abnormal.

$$PV+ = P(D+ | T+)$$

Negative predictive value is the probability that a patient with normal test results is truly normal.

$$PV- = P(D- | T-)$$

Using the same data as above, we know that 231 of 263 patients with abnormal liver scans had abnormal pathology, giving the proportion of correct diagnoses as

$\hat{P}V+ = 231/263 = 0.88$ . Similarly, among the 81 patients with normal liver scans the proportion of correct diagnoses was  $\hat{P}V- = 54/81 = 0.59$ . 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.

$$p = \text{Prevalence} = P(D+)$$

In the liver scan study, the estimated prevalence of the abnormality was  $\hat{p} = 0.75$ . If the same test was used in a different clinical setting where the prevalence of the abnormality was 0.25, we would have an estimated positive predictive value of 0.45 and an estimated negative predictive value of 0.95. [ see formulae below ] 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 (PV+ and PV-) can be determined for any prevalence by using Bayes Rule and getting:

$$PV+ = \frac{p Sn}{p Sn + (1 - p)(1 - Sp)}$$

$$PV- = \frac{(1 - p)Sp}{(1 - p)Sp + p(1 - Sn)}$$

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 values 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 [LR], calculated as sensitivity/ (1 - specificity).

$$LR = \frac{Sn}{1 - Sp}$$

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 estimated to be 0.75, so the pre-test odds of disease was estimated as  $0.75/(1 - 0.75) = 3.0$ . The sensitivity was estimated as 0.895 and the specificity was 0.628. The post-test odds of disease given a positive test is  $0.895/(1 - 0.628) = 2.41$ , 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.

$$\frac{PV+}{1 - PV+} = LR \frac{p}{1 - p}$$

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.

In Stata, you can download sbe36.1 and then -

. diagti 231 27 32 54

True disease status	Test result		Total
	Pos.	Neg.	
Abnormal	231	27	258
Normal	32	54	86
Total	263	81	344

Sensitivity	Pr ( +  D)	89.53%	85.14%	92.99%
Specificity	Pr ( -  ~D)	62.79%	51.70%	72.98%
Positive predictive value	Pr ( D  +)	87.83%	83.26%	91.53%
Negative predictive value	Pr (~D  -)	66.67%	55.32%	76.76%
Prevalence	Pr (D)	75.00%	70.08%	79.49%

. diagti 231 27 32 54,prev(50)

True disease status	Test result		Total
	Pos.	Neg.	
Abnormal	231	27	258
Normal	32	54	86
Total	263	81	344

Sensitivity	Pr ( +  D)	89.53%	85.14%	92.99%
Specificity	Pr ( -  ~D)	62.79%	51.70%	72.98%
Positive predictive value	Pr ( D  +)	59.65%	.%	.%
Negative predictive value	Pr (~D  -)	41.00%	.%	.%
Prevalence	Pr (D)	50.00%	.%	.%