Sensitivity and Specificity

Contents

Purpose of measuring sensitivity and specificity

"The simplest diagnostic test is one where the results of an investigation…..are used to classify patients into two groups according to the presence or absence of a symptom or sign"

Altman and Bland (1994)

However, there is a lot of variability between humans, and what is 'normal' for you may be abnormal for me, so it is very rare for a diagnostic test to be correct all the time, in all people. Sometimes a person will have a 'positive' result from a diagnostic test, but further investigations show them to be disease free. This is known as a 'false positive'. On other occasions a test may come back clear yet the person actually has the condition – this is known as a 'false negative'.

Sensitivity and specificity are two measures used to evaluate the diagnostic accuracy of a test, by looking at the proportion of correct diagnosis that a test gives.

- Sensitivity: The ability to correctly diagnose disease.
- Specificity: The ability to discriminate healthy from diseased patients

Calculating sensitivity and specificity

When evaluating a diagnostic test you need something to compare it to: a 'Gold standard', which is the definitive diagnose. If such a standard exists why do we need another test, you ask? Good question: the Gold Standard may involve invasive, costly or dangerous tests, or require waiting till the condition has become more advanced or the patient has died.

Ok, let's consider an example from Altman and Bland (1994) where the result of a diagnostic liver scan is compared with the definitive diagnosis obtained later (either after death, from biopsy or during surgery).

	Definitive diagnosis		
Diagnostic scan	Abnormal (+)	Normal (-)	Total
Abnormal (+)	231	32	263
Normal (-)	27	54	81
Total	258	86	344

From this we can see that of the 258 patients who actually had liver disease, this was detected by the diagnostic scan in 231 cases: i.e. 231/258*100=89.5%. This is the sensitivity of the test.

However the scan only gave 54/86 (62.8%) of patients the 'all clear'. This is the Specificity of the test.

Putting it more generally:

	Definitive diagnosis		
Diagnostic test	Abnormal (+)	Normal (-)	Total
Abnormal (+)	**True positive**	False positive	
Normal (-)	False negative	*True negative*	
Total	**Total with condition**	*Total without condition*	

Sensitivity = **True positive**/**Total with condition** * 100 %

Specificity = *True negative*/*Total without condition* *100%

Ideally both Sensitivity and Specificity will be 100%

Consequences

Consider two diagnostic tests:

“Echocardiography had a sensitivity of 93% and a specificity of 95% in diagnosing right ventricular hypertrophy (RVH)...

electrocardiography (ECG) had a sensitivity of 31% and a specificity of 85% in diagnosing RVH. “

Prakash 1981

Which, echocardiology or electrocardiology, is the better test – in terms of sensitivity and specificity?

Echocardiology, as it is closer to 100% for both sensitivity and specificity, is clearly the superior test in this case. However, it is often not that clear cut, frequently there is a trade-off between sensitivity and specificity. The implications of poor sensitivity and/or specificity vary according to the clinical situation, Consider some extreme examples, for each one decide which is the more critical parameter, Sensitivity or Specificity:

1. A life-threatening condition which can be successfully treated easily, cheaply and with few side effects, but only if it is diagnosed early

2. A slow growing cancer, removal requires major surgery, symptoms are very similar to an uncomfortable condition which usually clears up on it is own after a few weeks.

References:

Altman D G, Bland J M. 1994, Statistics Notes: Diagnostic tests 1: sensitivity and specificity BMJ 1994;308:1552

Kirkwood &Sterne 2003, Essential medical statistics, 2nd ed., Blackwell science, Oxford Pp.430-433.

Prakash, R; 1981 Echocardiographic diagnosis of right ventricular hypertrophy: Correlation with ECG and necropsy findings in 248 patients, Catheterization and Cardiovascular Diagnosis 1981 7:2:179-184.