THE UNIVERSITY of EDINBURGH | usher institute

Statistics of Diagnostic Tests

Basic Principles

Niall Anderson

Supported by

THE UNIVERSITY of EDINBURGH | DDI Data-Driven Innovation

Diagnostic Tests

Often need to use a “test” to classify samples/ individuals as 1 of 2 categories:

- Biologically active/ Inactive
- Disease/ No disease … etc

Test = measurement (categorical,…, continuous)

Often need threshold for test results

e.g. DBP > 90 mmHg => Hypertension

Haemoglobin < 12 g/dl => Anaemia

If test is to be useful, need to know how reliable it is.

What is “reliable” in this context?

Sample/ patient may or may not have the characteristic of interest

Test gives a result indicating one of 2 possible outcomes

How do these match?

Look at some ways of measuring test performance…

Notation/ Basics

D+	“Disease”/ Outcome present
D-	“Disease”/ Outcome absent
T+	Test positive
T-	Test negative

Pr (A | B) = probability that A happens, *given that* B happens

Need “gold standard” diagnosis to get D+/-

Example

Exercise test to diagnose CAD.

Positive test defined by changes on EEG.

Disease diagnosed definitively by angiogram.

		Coronary Artery Disease		Total
		Present (D+)	**Absent (D-)**	
Exercise Tolerance Test	**Positive (T+)**	815	115	930
	Negative (T-)	208	327	535
	Total	1023	442	1465

		Coronary Artery Disease		Total
		Present (D+)	Absent (D-)	
Exercise Tolerance Test	Positive (T+)	815	115	930
	Negative (T-)	208	327	535
	Total	1023	442	1465

Define two measures of test performance

Sensitivity: Pr(T+ | D+) = 815/1023 = 0.80
(Tells us about the test in a positive sense)

Specificity: Pr(T- | D-) = 327/442 = 0.74
(Gives information about negative performance)

Good test would have high sensitivity **and** specificity (close to 1)

Prevalence : Pr(D+) = 1023/1465 = 0.70
i.e. estimate of frequency of outcome in population

Sensitivity and Specificity are **independent** of prevalence

(try doubling number of D+ people in each cell, then recalculating sensitivity and specificity).

Useful property – measures are comparable between populations for same test.

May also see **false positive rate** = 1 - specificity
false negative rate = 1 - sensitivity