Nemours Biostatistics Core Statistics Course

Li Xie

March 5, 2014

Outline

- An Instance of Hypothesis Testing: Medical Diagnostic Testing
- Relating Back to the Basics: Sampling, Central Tendency Measures (Mean, Median), Variability Measures (Variance, Standard Deviation)

Pt encounter

- Symptoms (relevant & irrelevant), Health history, etc.

Medical tests

- Exclusionary
- Confirmatory
- Single differential ddx
- Multiple differential ddx

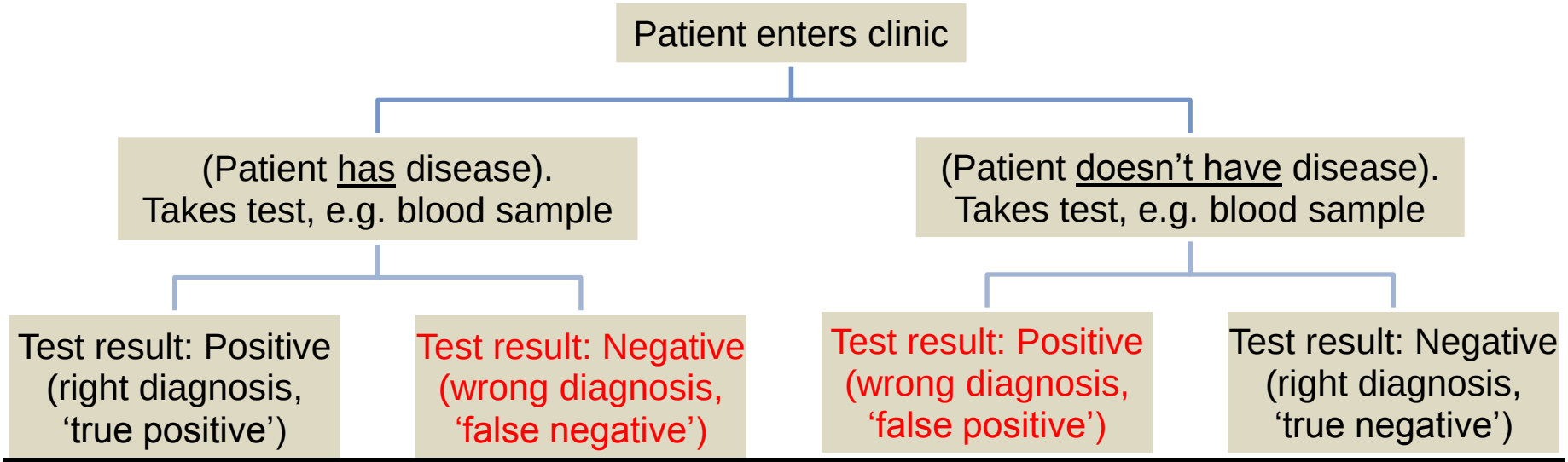
Diagnosis

Sample of all possible data of a dx (relevant & irrelevant)

Hypothesis tests

- Reject some hypotheses
- Accept some hypotheses
- Simple hypothesis
- Composite hypothesis

Statistical inference



		'TRUTH' (by Reference test – Gold standard)		Total
		+	-	
TEST RESULT (by Index test)	+	True Positive	False Positive	Total said to have dx
	-	False Negative	True Negative	Total said not to have dx
Total		Total with dx	Total without dx	

		'TRUTH' (by Reference test – Gold standard)	
		+	-
TEST RESULT (by Index test)	+	True Positive	False Positive
	-	False Negative	True Negative

	H_0 is false (yes dx)	H_0 is true (no dx)
Reject H_0 (test +)	Right decision	Wrong decision Type I Error
Accept H_0 (test -)	Wrong decision Type II Error	Right decision

Terminology

- Validity – test's ability to indicate which individuals have the disease and which do not; in terms of sensitivity & specificity
- Reliability – repeatability (on the same pt, btw pts, btw raters)
- Yield - # of tests that can be done in a time period
- Sensitivity – test's ability to identify correctly those who have the dx
- Specificity – test's ability to identify correctly those who are dx-free
- The predictive value of a positive test (PPV) is the probability that an individual testing positive is truly affected
- The predictive value of a negative test (NPV) is the probability that an individual testing negative is truly non-affected
- Sensitivity & specificity are **unconditional** characteristics of the test
- Predictive values are **conditional** (on dx prevalence) characteristics

The 2 X 2 table

		'TRUTH'	
		+	-
TEST	+	True Pos (a)	False Pos (b)
	-	False Neg (c)	True Neg (d)

Formulas:

Sensitivity = a / a+c

Specificity = d / b+d

Accuracy = a+d / a+b+c+d

Prevalence = a+c / a+b+c+d

Predictive Value:

positive = a / a+b

negative = d / c+d

Positive Test = a+b / a+b+c+d

Negative Test = c+d / a+b+c+d

Diseased = a+c / a+b+c+d

Not Diseased = b+d / a+b+c+d

*** a,b,c,d are counts of pts**

Tests based on continuous variables

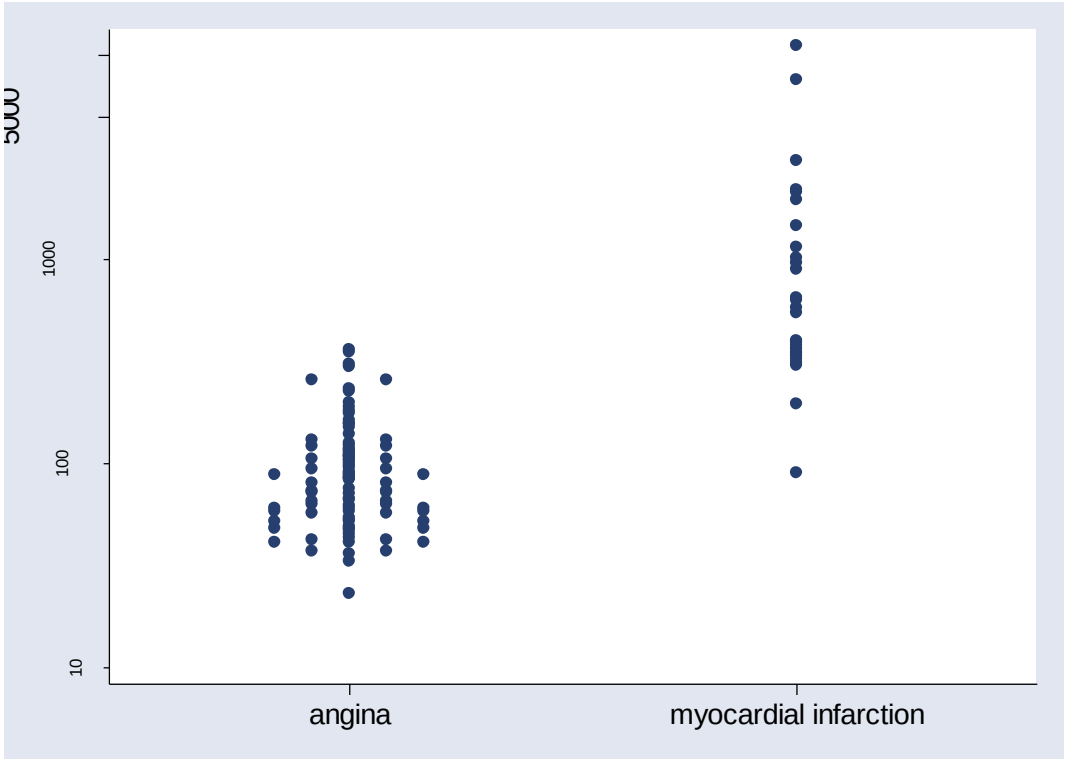
- One or more continuous variables can be a marker for a condition, where a very low/high level indicates a low/high likelihood of having the condition.

- A cut-off level can be determined where having higher/lower than that cut-off indicates a positive test result.

- Different cut-off points will give different sensitivity/specificity values.

Tests based on continuous variables

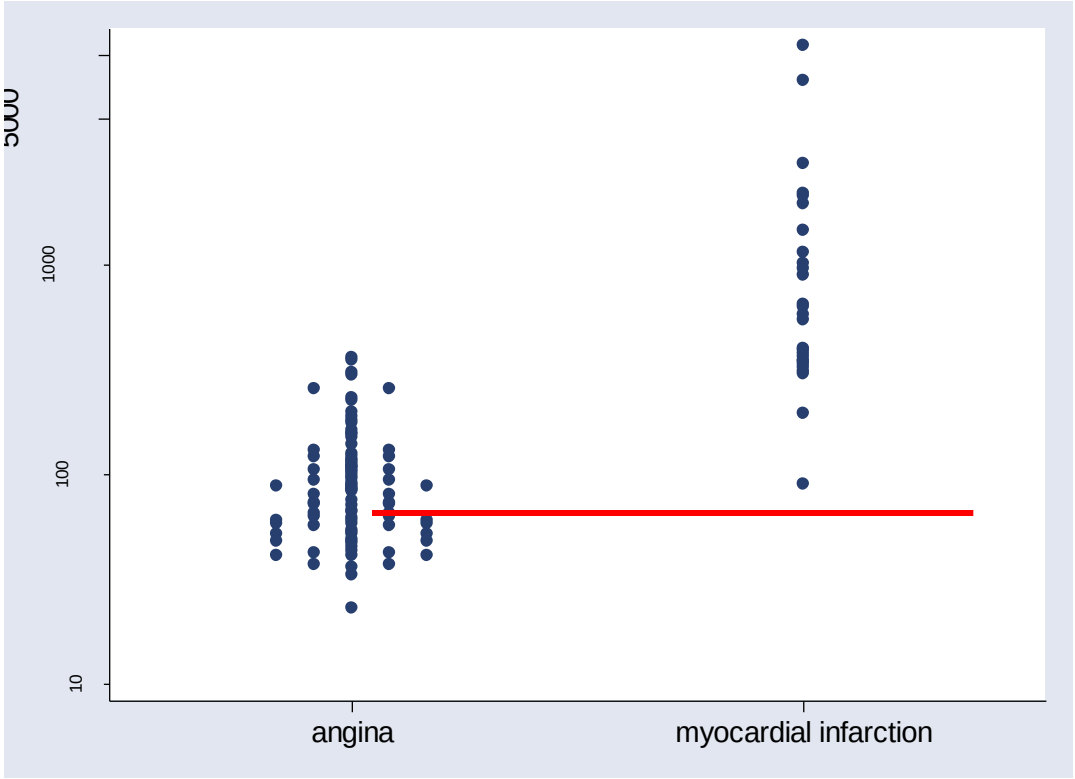
E.g. Creatine kinase (CK) in patients with unstable angina or acute myocardial infarction



Data of Frances Boa, from 'An introduction to Medical Statistics' by Martin Bland

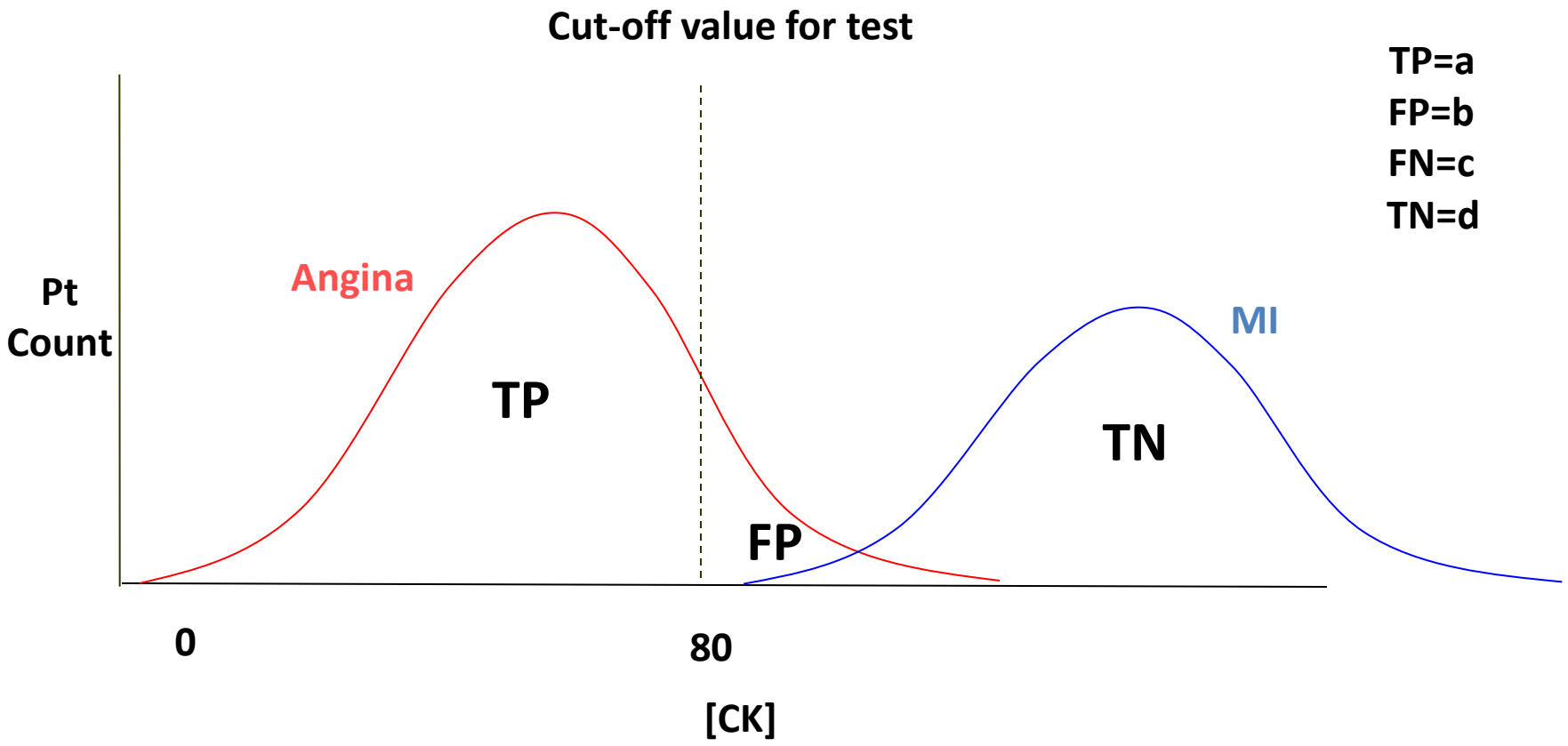
Tests based on continuous variables

		MI Status	
		+	-
[CK] test	+ (≥80)	27	54
	- (<80)	0	39
Total		27	93



Cut-off level at 80

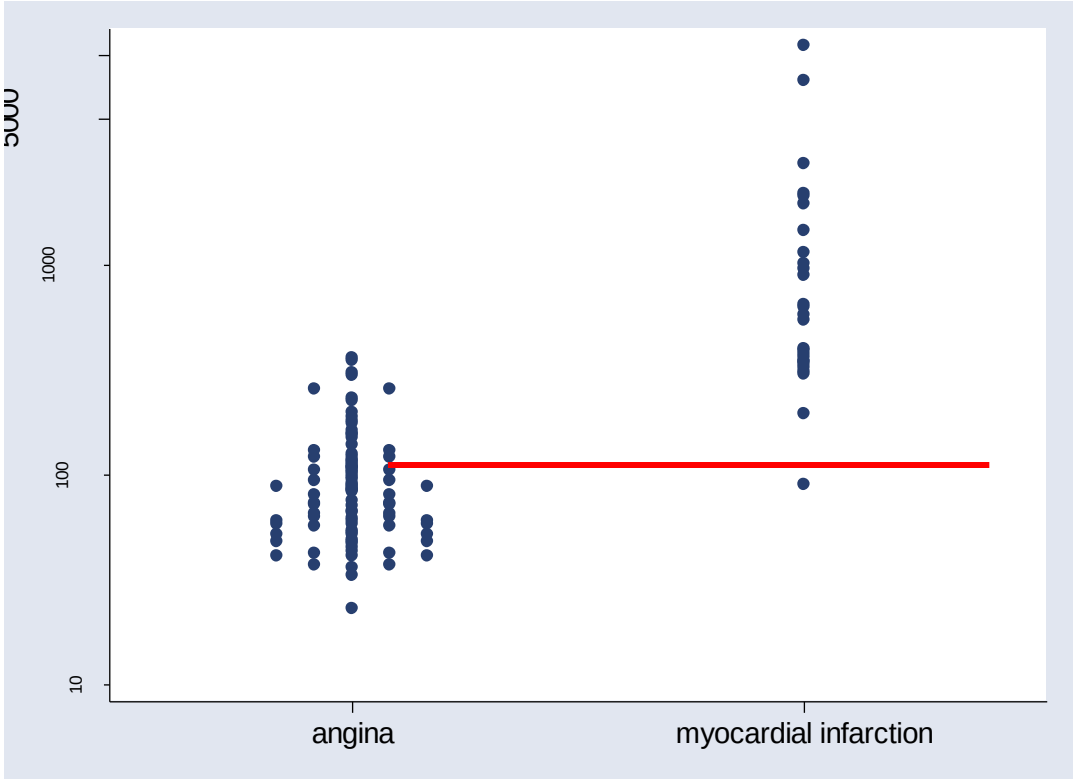
Tests based on continuous variables



*There is no FN in this example

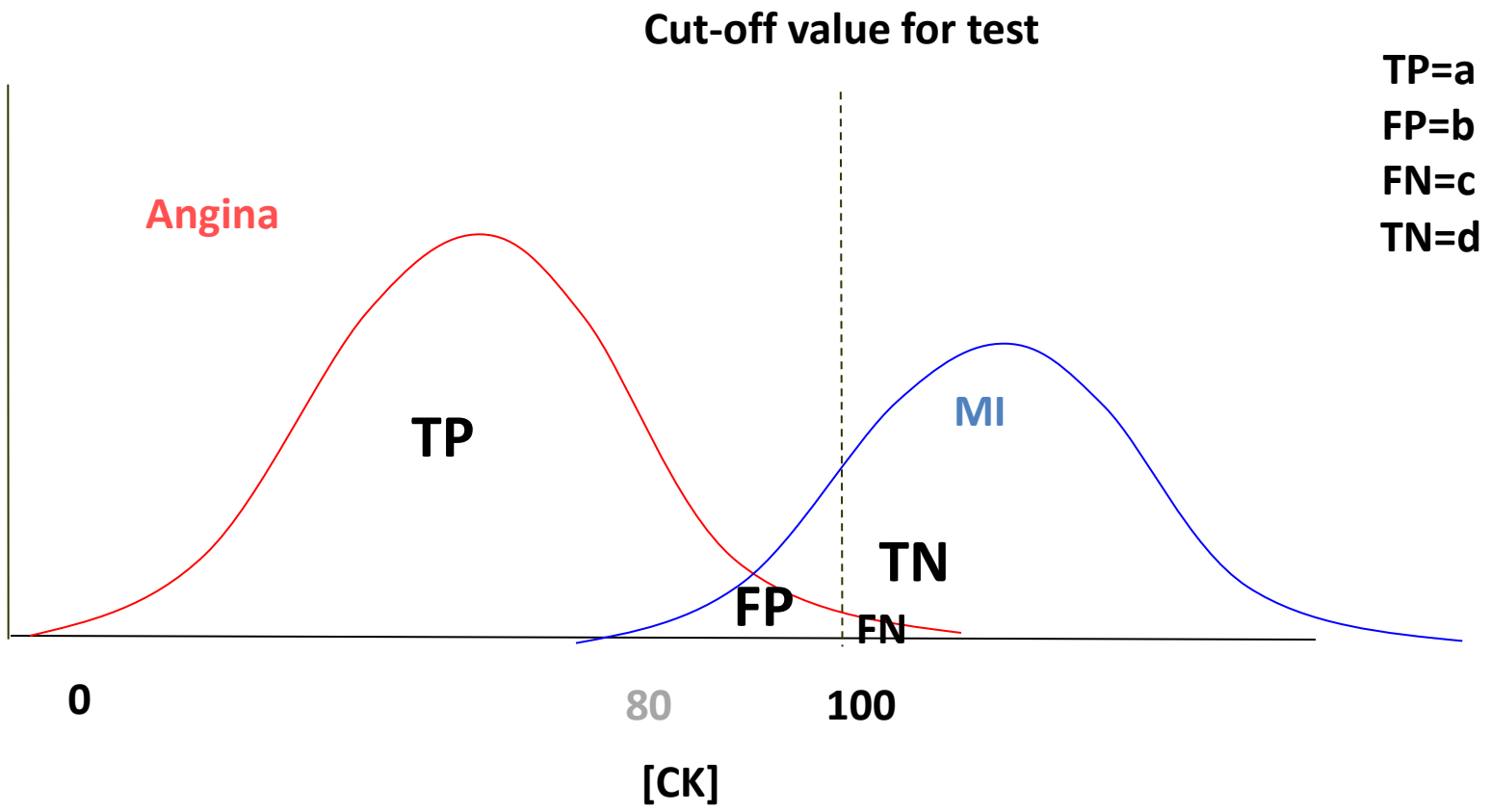
Tests based on continuous variables

		MI Status	
		+	-
[CK] test	+ (≥100)	26	35
	- (<100)	1	58
Total		27	93



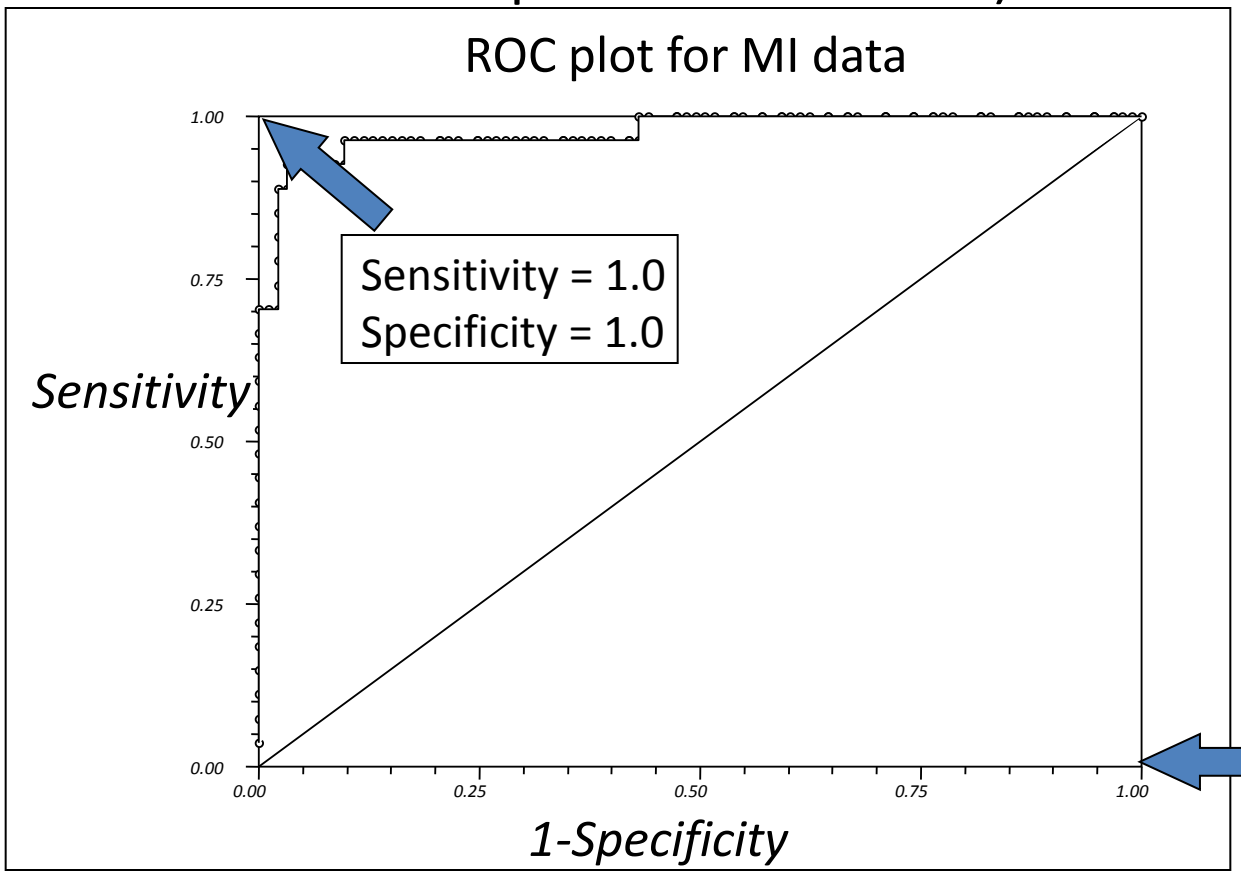
Cut-off level at 100

Tests based on continuous variables



The trade-off

- Plot sensitivity against (1-specificity) to get the ROC ('receiver operating characteristic') curve.
- Ideally want high sensitivity and high specificity (but increase in one is at expense of the other).



The diagonal line represents

sensitivity = specificity

i.e. taking the test is as good as flipping a coin.

Sensitivity = 0.0
Specificity = 0.0

Optimum cut-off

MI data:

- ‘Optimum’ cut-off point selected = 302
- Sensitivity (95% CI) = 0.93 (0.76 to 0.99)
- Specificity (95% CI) = 0.97 (0.91 to 0.99)

Note: “Optimum” assumes sensitivity and specificity of equal concern. In some cases, “Optimum” may also involve cost-effectiveness and other non-clinical considerations.

Strategic choices of cut-off points for continuous results

Consider the implications of the two possible errors:

- If *false-positive* results must be avoided (such as the test result being used to determine whether a patient undergoes dangerous surgery), then the cutoff point might be set to maximize the test's *specificity*
- If *false-negative* results must be avoided (as with screening for neonatal phenylketonuria), then the cutoff should be set to maximize the test's *sensitivity*

Area under the ROC curve

- Area under the ROC curve can be between 0 (sensitivity and specificity always 0.0) and 1 (sensitivity and specificity always 1.0).
- Can be useful for comparing two tests.
- MI data: Area under curve is an estimate of ‘probability that CK of random person with MI will be higher than for random person with angina’.

Comparative ROC Curves