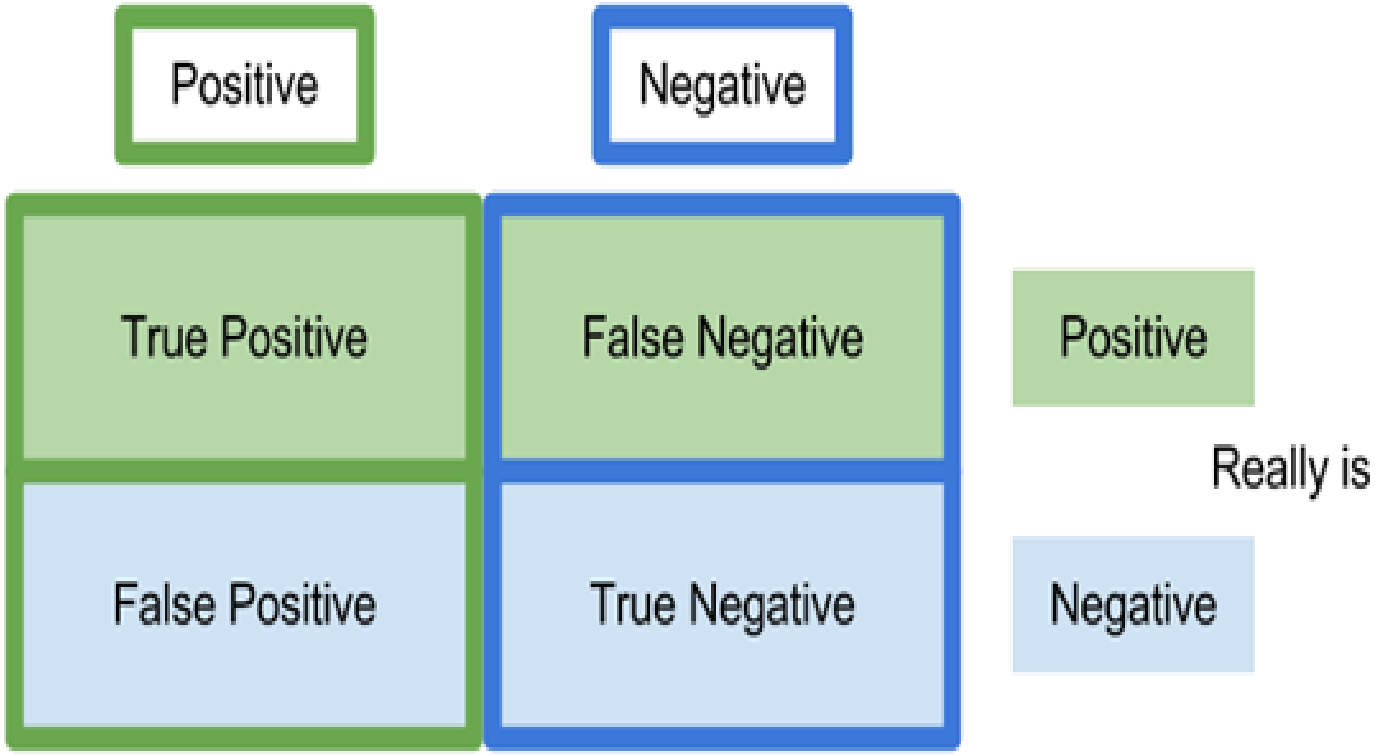
Topic Number Six

Evaluation of diagnostic tests



Kamel Adwan, Ph.D
Department of Biology and Biotechnology
An Najah N. University

Reliability

- Reliability can be divided in reproducibility and accuracy.

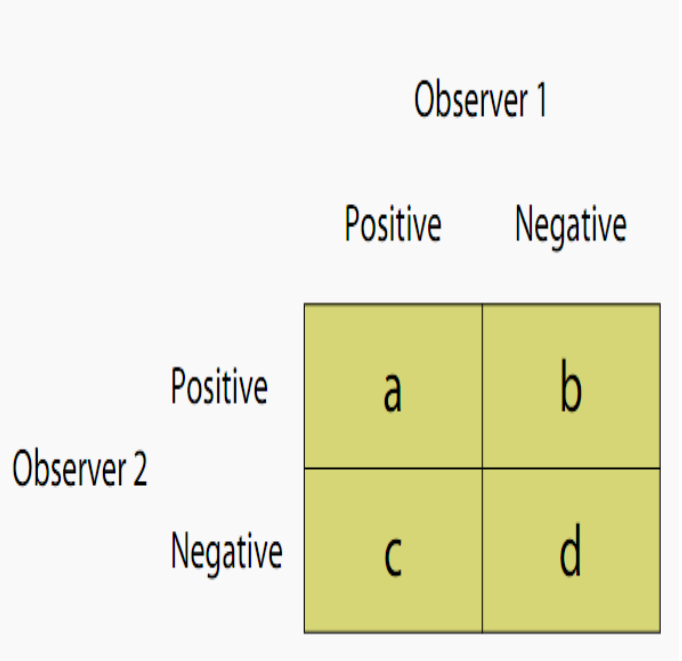
Reproducibility, also called is precision

- Intra-subject variation is a variation in the results of a test conducted on two separate occasions, made by one observer (intra = within). The difference is due to the changes (such as physiological, environmental, etc.)
- Inter-observer variationis a variation in the result of a test due to multiple observers examining the result (inter = between). The difference is due to the variation in the selection of an appropriate cut-off point to separate 'positive' and 'negative' results

Accuracy, also called validity

- The level of agreement between the test result and the "true" clinical state i.e., which individuals have the disease and which do no

- When a diagnostic test has to be evaluated, a gold standard test as reference is needed. The "gold standard" is a diagnostic test that correctly classify diseased and non-diseased animals.
- A theoretical example of the results of a reproducibility study of two observers A and B is shown.
- The squares with a and b represent the number of patients with positive tests as judged positive by observer A.
- The squares with a and c represent the number of patients with positive tests as judged by observer B.
- The squares a, b and c represents the number of patients with positive tests as judged by either one or both observers among the total patients n.



Percent Agreement

$$\text{Overall Percent Agreement} = \frac{a + d}{a + b + c + d} \times 100$$

$$\text{Percent Positive Agreement} = \frac{a}{a + b + c} \times 100$$

Note: This is a conditional probability

Example

		Radiologist A: Positive	Radiologist A: Negative
Radiologist B	Positive	4	5
	Negative	2	6

$$\text{Overall Percent Agreement} = \frac{4 + 6}{4 + 5 + 2 + 6} \times 100 = 58.8\%$$

$$\text{Percent Positive Agreement} = \frac{4}{4 + 5 + 2} \times 100 = 36.4\%$$

Sensitivity and Specificity

Sensitivity

The ability of the test to identify correctly those who have the disease (i.e. proportion of true positives).

Specificity

The ability of the test to identify correctly those who do not have the disease (i.e. proportion of true negatives).

		Disease +	Disease −
Test	+	**a** (True positives)	**b** (False positives)
	−	**c** (False negatives)	**d** (True negatives)

Determining the Sensitivity, Specificity of a New Test

- Use a 2 x 2 table to the right to compare the performance of the new test to the gold standard test
- Sensitivity is the ability of the test to identify those who have the disease by the new test from all individuals with the disease
- Specificity is the ability of the test to identify those who do not have the disease by the new test from all individuals free from the disease

		Truth or Gold std test: +	Truth or Gold std test: −
Your test	+	**a** (True positives)	**b** (False positives)
	−	**c** (False negatives)	**d** (True negatives)

$$\text{sensitivity} = \frac{a}{a+c} = \frac{\text{true positives}}{\text{disease}+}$$

$$\text{specificity} = \frac{d}{b+d} = \frac{\text{true negatives}}{\text{disease}-}$$

Calculating Sensitivity and Specificity

- Assume a population of 1,000 people
- 100 have a disease
- 900 do not have the disease
- Results of the screening appears in the table to the right

Screening Results	True Characteristics in Population		Total
	Disease	No Disease	
Positive	80	100	180
Negative	20	800	820
Total	100	900	1,000

Sensitivity = 80/100 = 80%

Specificity = 800/900 = 89%

Positive and Negative Predictive Value

Positive predictive value (PPV)

The proportion of patients who test positive who actually have the disease (true positives)

Negative predictive value (NPV)

The proportion of patients who test negative who are actually free of the disease (true negatives)

		Disease +	Disease −
Test	+	**a** (True positives)	**b** (False positives)
	−	**c** (False negatives)	**d** (True negatives)

$$\text{Positive predictive value} = \frac{a}{a+b} = \frac{\text{True Positives}}{\text{Test}+}$$

$$\text{Negative predictive value} = \frac{d}{c+d} = \frac{\text{True Negatives}}{\text{Test}-}$$

Calculating Predictive Values

- Assume a population of 1,000 people
- 100 have a disease
- 900 do not have the disease
- Results of the screening appears in the table to the right

Positive predictive value = 80/180 = 44%

Screening Results	True Characteristics in Population		Total
	Disease	No Disease	
Positive	80	100	180
Negative	20	800	820
Total	100	900	1,000

Negative predictive value = 800/820 = 98%

True prevalence estimation

True prevalence is the proportion of animals in the population which really do have the disease regardless of their test result. it includes the “true” positives and the “false” negatives (a+c/ n).

Apparent Prevalence is the proportion of animals in the population giving a positive test result regardless of their true status for the disease. It includes the “true” positives and the “false” negatives. (a +b/n).

If you know the sensitivity and specificity of a test, you can calculate the apparent prevalence to the true prevalence using the formula to the right

		Truth or Gold std test +	Truth or Gold std test −
Your test	+	**a** (True positives)	**b** (False positives)
	−	**c** (False negatives)	**d** (True negatives)

n = a+ b+ c+ d

$$TPr = \frac{APr + Sp - 1}{Se + Sp - 1}$$

Predictive values and Prevalence

As prevalence increases, positive predictive value (PPV) increases and negative predictive value (NPV) decreases.

Example: Sensitivity = 99%; Specificity = 95%					
Disease Prevalence	Test Results	Sick	Not Sick	Totals	Positive Predictive Value
1%	+	99	495	594	$\frac{99}{594} = 17\%$
	–	1	9,405	9,406	
	Totals	100	9,900	10,000	
5%	+	495	475	970	$\frac{495}{970} = 51\%$
	–	5	9,025	9,303	
	Totals	500	9,500	10,000	

PPV Formula

Formulae for calculating predictive values are based on Bayes' theorem and are as follows

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

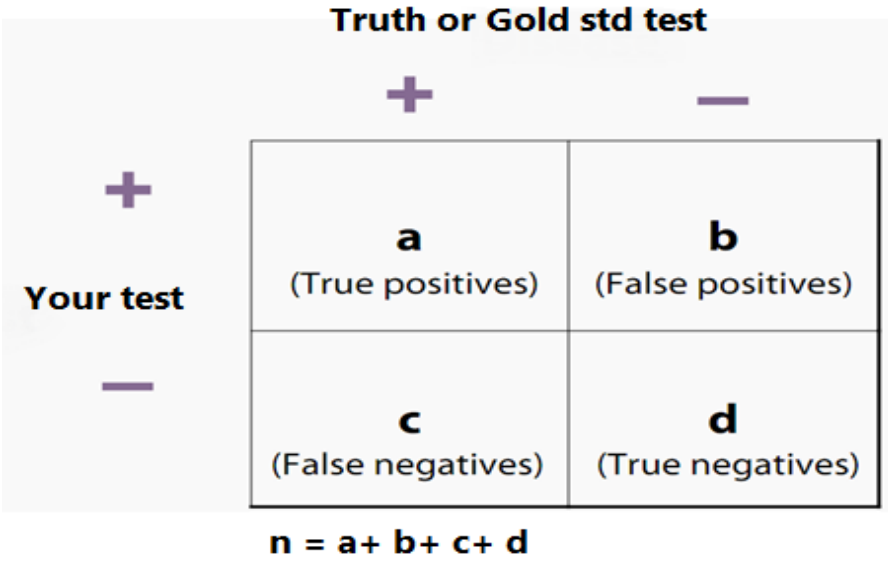
$$NPV = \frac{\text{specificity} \times (1 - \text{prevalence})}{[(\text{specificity} \times (1 - \text{prevalence})] + [(1 - \text{sensitivity}) \times \text{prevalence}]}$$

kappa statistic

- If no 'gold standard' exists for a particular condition and it is necessary to evaluate the diagnosis between different tests, the kappa test can be used to measure the level of agreement.
- The kappa statistic lies within a range between –1 and +1
- The observed agreement given by the formula: OA = (a + d)/(a + b + c + d)
- The expected agreement which would be obtained is given by the formula: EA = [{(a + b)/n} x {(a + c)/n}] + [{(c + d)/n} x {(b + d)/n}]
- The maximum possible agreement beyond chance = (OA – EA) / (1-EA)

Evaluation of kappa statistic

Kappa value	Evaluation
> 0.81	Almost perfect agreement
0.61 – 0.80	Substantial agreement
0.41 – 0.60	Moderate agreement
0.21 – 0.40	Fair agreement
0.01 – 0.20	Slight agreement
0.00	Poor agreement



Example on kappa statistic

For example, using the data appeared in the table to the right.

OA = (74 + 127)/ 258 = 0.779

EA = [{129/258} x {76/258}] + [{129/258} x {182/258}]

= (0.500 x 0.295) + (0.500 x 0.705)

= 0.1475 + 0.3525

= 0.500

The maximum possible agreement beyond chance = 1 – 0.500 = 0.500

k = (0.779 – 0.5)/0.5

= 0.279/0.5

= 0.558 indicating moderate agreement between the two tests.

Note: The kappa value gives no indication which of the tests is better and that a good. The agreement may indicate that both tests are equally good or equally bad.

	Histology		
	+ve	-ve	
PCR +ve	74	55	129
PCR –ve	2	127	129
	76	182	258

Review

Review of Testing Terms and Formulas.

		DISEASE Status D +	DISEASE Status D –	
TEST Results	T +	a	b	a + b
	T –	c	d	c + d
		a + c	b + d	n

Apparent Prevalence	(a+b) / n
True Prevalence	(a+c) / n
Predictive Value of Positive test	a /(a + b)
Predictive Value of Negative test	d /(c + d)
Sensitivity	a /(a + c)
Specificity	d /(b + d)