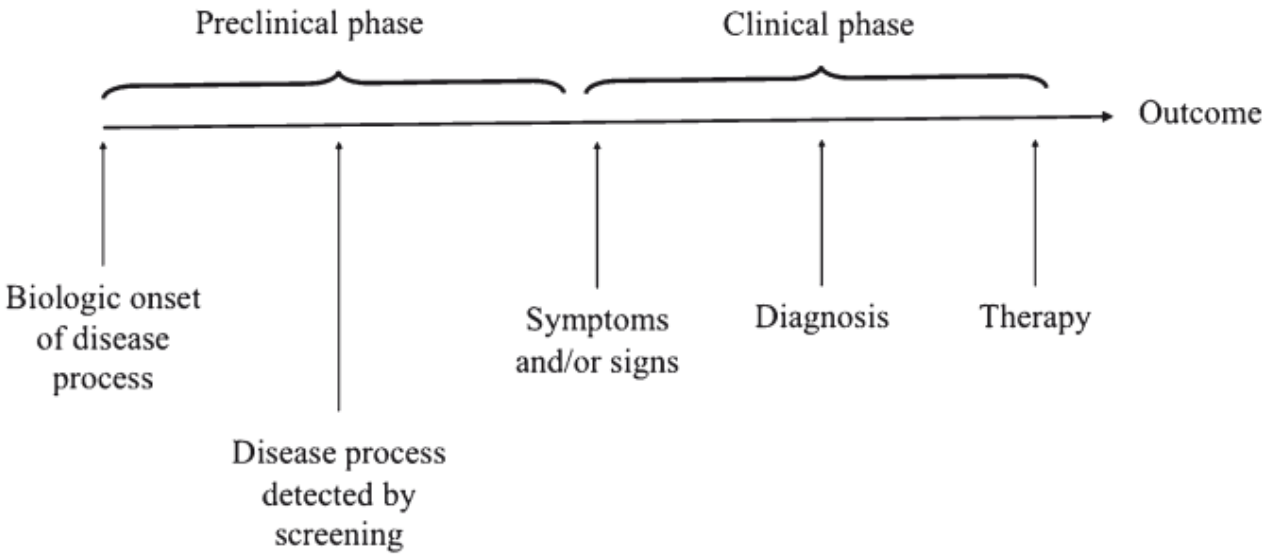
Screening


Preclinical phase
Clinical phase
Outcome
Biologic onset of disease process
Disease process detected by screening
Symptoms and/or signs
Diagnosis
Therapy

General Principles of Screening

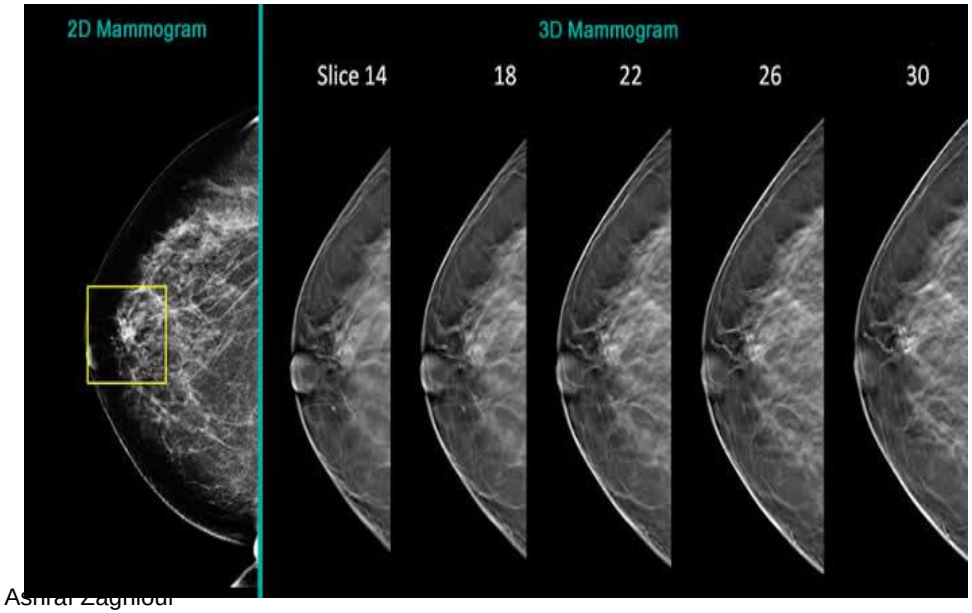
- Epidemiologic principles are used to gain insight into the causes of disease

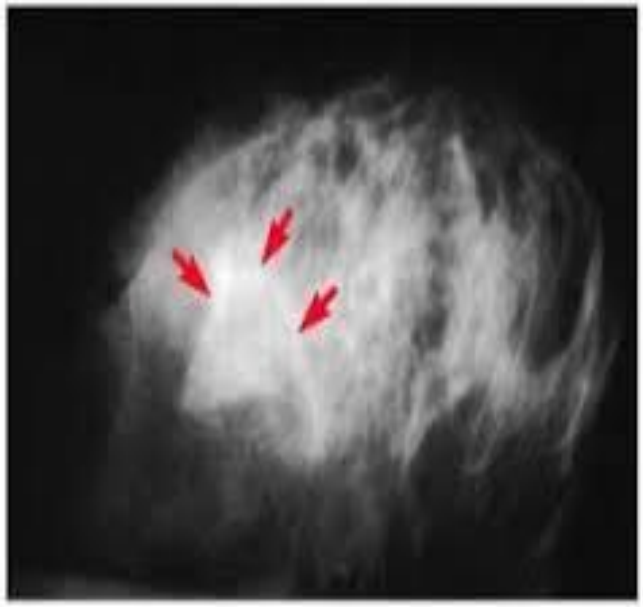
- Screening and diagnostic testing as clinical tools for **detecting** and **treating disease**

Screening refers to the early detection of a disease or condition in the preclinical phase, defined as the **period before clinical symptoms or signs are present**

Example

- The identification of previously unrecognized disease can lead to subsequent interventions that impact the course of the disease, for example screening mammography can detect breast cancer at an early stage before it is clinically apparent and surgery plus chemotherapy given at an early stage can cure the disease.


2D Mammogram
3D Mammogram
Slice 14
18
22
26
30



Ideally, screening tests are applied to clinical conditions that progress in a series of ordered steps

Characteristics of the **PERSON** being Screened

Appears to be free of the disease of interest

Not seeking care because they are not sick

Persuaded to be screened by the health service

Qualities of Diseases Appropriate for Screening

1. **The Disease should be Important in the Screened Population**

- **Generally, screening tests focus on serious diseases.** Examples include screening for colon cancer in middle-aged adults, and screening for phenylketonuria in newborns.

- Detection of these potentially fatal conditions can lead to interventions that dramatically reduce mortality.

Cancer Prostate in Men

- It is probable that many men in whom a cancer could be detected by screening (e.g. with a prostate specific antigen **(PSA)** test would never develop symptoms or suffer from the disease

2. Early Recognition and Treatment of the Disease Should Prevent Clinical Outcomes

- Detecting untreatable conditions earlier in their course can increase patient anxiety without influencing the disease process.

- For example, electron beam computed tomography (EBCT) is a specialized scanning procedure that is used to detect asymptomatic coronary artery disease. The EBCT scan can **rapidly quantify the extent of coronary artery calcification**, a marker of atherosclerosis.

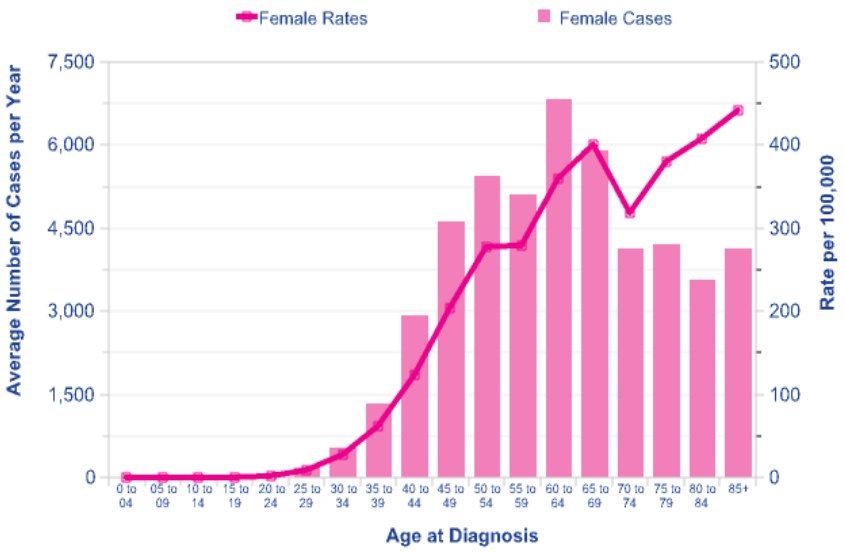
However, EBCT generally cannot distinguish high-grade coronary lesions that require surgical intervention from diffuse low-grade atherosclerotic plaques

3. The Disease Should have a Preclinical Phase

- It would be difficult to screen for a condition like the common cold, because the time from biologic onset of disease to clinical symptoms is so short.

- On the other hand, other diseases, such as colon cancer, have an ordered preclinical phase that can be detected by the presence of histologic findings or specialized radiographic imaging studies, or specific biomarkers.

4. The **PREVALENCE** of pre-clinical disease **should be high** among the population screened



Rate of New Prostate Cancer Cases by Race and Ethnicity: United States	
Race or Ethnicity	**Incidence Rate per 100,000**
All Races	156.9
White	145.1
Black	226.0
Asian/Pacific Islander	78.2
American Indian/ Alaska Native	71.7
Hispanic	121.6

Possible conditions for which screening considered

- **Infectious diseases**
  - Tuberculosis, HIV
- **Metabolic diseases**
  - Hypothyroidism in infancy, Phenylketonuria
- **Abnormalities in pregnancy**
  - Neural tube defects, Down's syndrome
- **Cancers**
  - Breast, prostate, cervical, colorectal
- **Cardiovascular diseases**
  - Hyperlipidaemia

Qualities of Screening Tests

1. General Qualities

- To achieve widespread use, a screening test ideally should be easy to administer, relatively inexpensive, and safe.

- Many blood tests and imaging studies satisfy these criteria, for example, the prenatal "**triple screen**" blood test that is used to screen for trisomy 21 during pregnancy and the **chest X-ray** that is used to screen for tuberculosis.

2. **Reliability and Validity**

- **Screening tests are generally judged by their reliability and validity**

- **Reliability** refers to the ability of a test to provide **consistent results**

- For example, the HIV antibody test is considered to be reliable, because the test will return a consistent result, positive or negative, within in a given individual on the same day.

On the other hand, the potassium hydroxide test for diagnosing cutaneous fungal infection may yield **different results** when repeated on the same individual, due to sampling variation, differences in specimen preparation, and the subjective opinion of the individual tester who is looking under the microscope

- **Validity** refers to the ability of a test to **detect true event (disease)**, as defined by some gold-standard measurement technique.

- For example, the validity of mammography for detecting breast cancer is typically judged against the gold standard diagnosis, which is made by breast biopsy and pathologic examination.

- The validity of serum creatinine levels for detecting kidney function is judged against formal measurement of the glomerular filtration rate performed using an intravenous tracer.

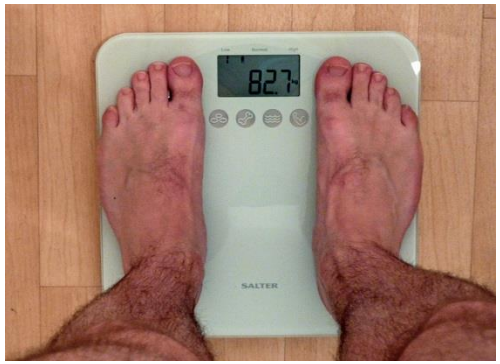
- Other clinical conditions are best diagnosed by expert opinion. For example, the gold standard diagnosis of heart failure in many clinical studies is considered to be the expert opinion of a panel of cardiologists, who review each patient's medical chart

In general, gold standard testing is invasive, expensive, or not practical to apply to a large population for screening purposes

Another perspective on reliability and validity is to think of an unreliable test as having random error **(changes in the environment and circumstances surrounding the study measurement)**

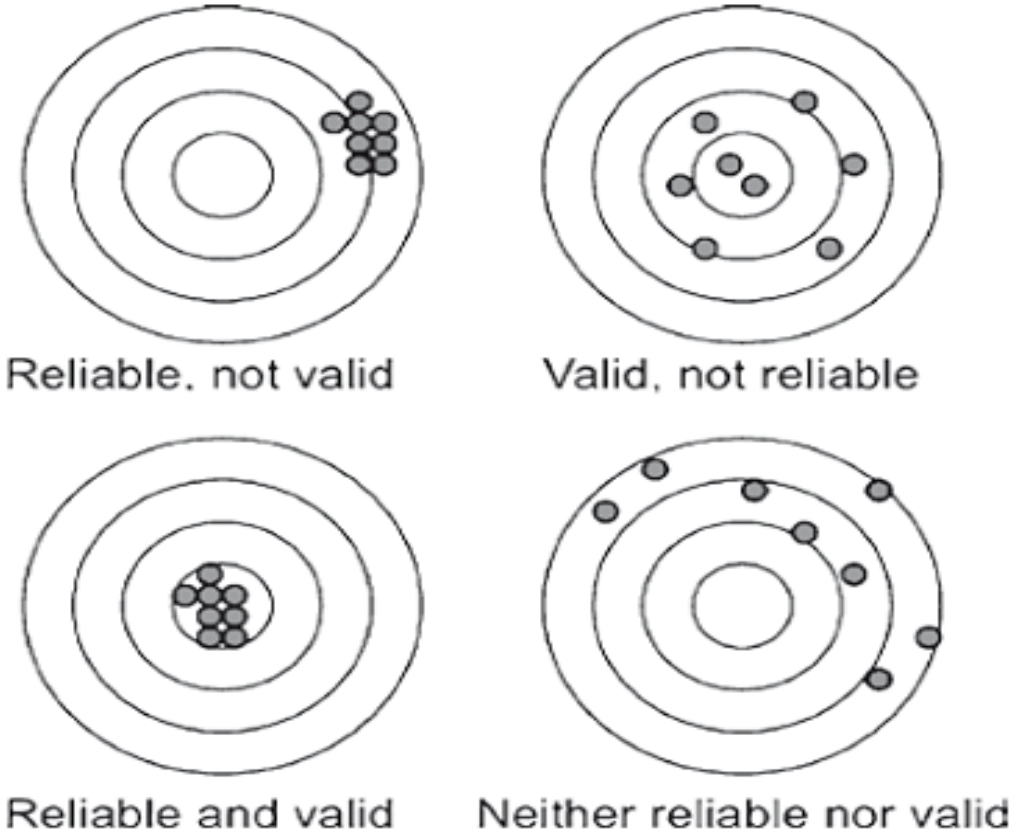
and

an invalid test as having systematic error **(changes in the instrument detecting the measurements)**


827
SALTER


testo 435
1037
9344


File Log Probe View
AdvancedSense
Environmental Test Meter
IQ-610


Reliable, not valid
Valid, not reliable
Reliable and valid
Neither reliable nor valid

Validity of Screening Tests

1. Sensitivity and Specificity

- **The validity of a screening test is typically described by the sensitivity and specificity of the test.**

- These terms describe how well the test performs compared to a gold standard test.

Sensitivity and specificity can be explicitly defined using a 2×2 table, in which true disease status is presented across the top of the table and test result status is presented on the left-hand side of the table

Sensitivity and specifcity of a screening test

	Disease	
	Yes	No
Test result		
Positive	*a*	*b*
Negative	*c*	*d*

$$\text{Sensitivity} = \frac{\text{Number who test positive with disease } (a)}{\text{Number with disease } (a+c)}$$

$$\text{Specificity} = \frac{\text{Number who test negative without disease } (d)}{\text{Number without disease } (b+d)}$$

- Sensitivity is the probability of testing positive given the presence of disease

- Specificity is the probability of testing negative given the absence of disease

Example

- For example, the **sensitivity** of mammography for detecting breast cancer among women over 50 years old is about 85%

- The interpretation of this sensitivity value is, "among women with biopsy proven breast cancer, the chance of having a positive mammogram is 85%"

Example

- The specificity of mammography for detecting breast cancer among women over 50 years old is about 95%.

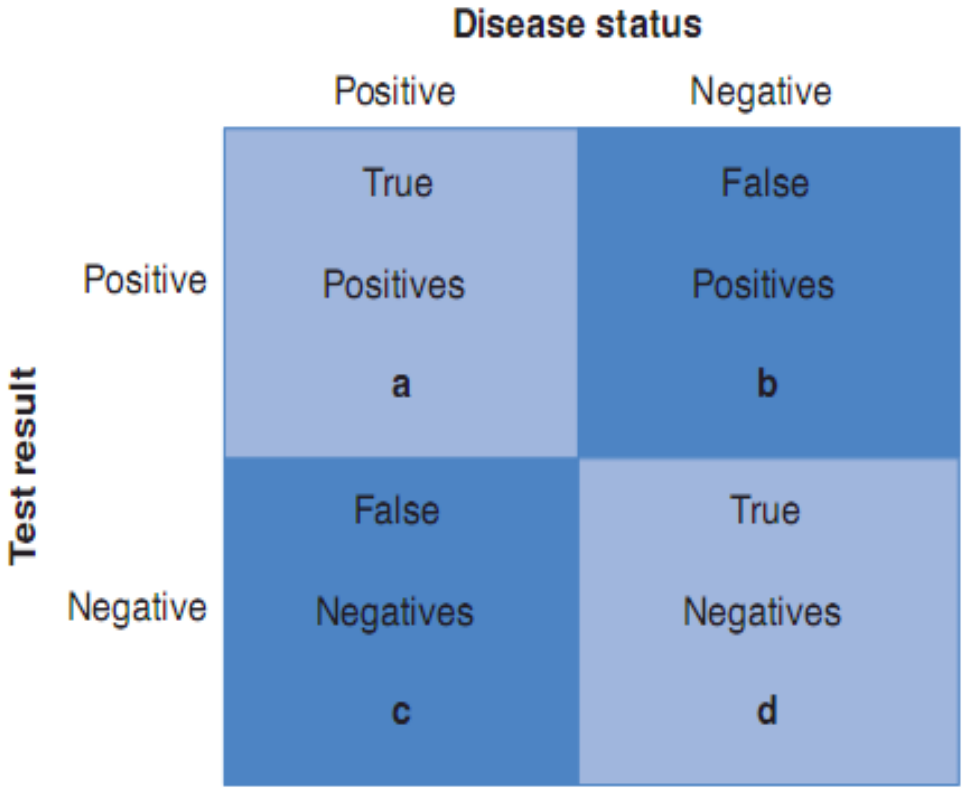
- The interpretation of this specificity value is, “among women with biopsy proven absence of breast cancer, the chance of having a negative mammogram is 95%.”

Example

- Children in many countries undergo a simple **hearing test** in their first year at school.

- Any who fail this screening test are retested at a later date and/or referred to a hearing clinic for further, more extensive tests to identify whether they have a real hearing problem.

Imagine that in a group of 500 children, 50 have a genuine hearing problem. Of these, 45 fail the school hearing test, as do 30 of the children with normal hearing (perhaps they had a cold on the day of the test)


Disease status
Positive
Negative
Test result
Positive
Negative
True
Positives
a
False
Positives
b
False
Negatives
c
True
Negatives
d

Calculate the sensitivity and specificity of the hearing test

	True hearing status		
School hearing test	Hearing problem	Normal	Total
Fail (positive test result)	45	30	75
Pass (negative test result)	5	420	425
Total	50	450	500

- Sensitivity and specificity characteristics generally remain consistent across different populations, or may vary to only a small degree.

- However, sensitivity and specificity characteristics do not provide important clinical information for individual patients. In the mammography example, women typically would not be interested in their probability of having a positive mammogram after they are diagnosed with breast cancer.

- Instead, they would like to know the opposite information, specifically, what is their chance of having breast cancer given a positive or negative mammography result? To answer this more clinically relevant question, two additional characteristics of screening tests are needed.

Positive and Negative Predictive Value

- **Positive predictive value** is the probability of disease given a positive test result. (how likely it is that a positive test result indicates the presence of disease. *It is the percentage of all people who test positive who really have the disease*)

$$\text{Positive predictive value} = \frac{\text{Number who test positive with disease } (a)}{\text{Number who test positive } (a+b)}$$

- Negative predictive value is the probability of no disease given a negative test result. (*the percentage of all people who test negative who really do not have the disease*)

$$\text{Negative predictive value} = \frac{\text{Number who test negative without disease } (d)}{\text{Number who test negative} (c + d)}$$

Calculate the PPV and the NPV for the hearing test

Example

- For example, the positive predictive value of mammography is 10% in a low-risk patient population.

- The interpretation of this positive predictive value is, “among low-risk women who have a positive mammogram, the probability of breast cancer is 10%.”

- The negative predictive value of mammography in this same population is 98%.
- The interpretation of this negative predictive value is, “among low-risk women with a negative mammogram, the probability of having breast cancer is 2%.”

- Positive and negative predictive values depend on the **prevalence of disease in the screened population.**

Example

The sensitivity and specificity of the hepatitis C antibody test for detecting hepatitis C infection are 99% and 95%, respectively.

What is the positive predictive value of the hepatitis C antibody test for detecting hepatitis C infection?

Hepatitis C antibody testing: disease prevalence unknown

	Disease		
	Yes	No	
Test result			
Positive	*a*	*b*	?
Negative	*c*	*d*	?
	Sensitivity = $a/(a+c) = 0.99$	Specificity = $d/(b+d) = 0.95$	

Positive predictive value = a / (a+b) = ?

- Given only the sensitivity and specificity characteristics of a test as in the table.
- It is not possible to determine the positive or negative predictive value
- More information is needed.

Example

The sensitivity and specificity of the hepatitis C antibody test for detecting hepatitis C infection are 99% and 95%, respectively. Among United States veterans, the prevalence of hepatitis C infection is 10%.

What is the positive predictive value of hepatitis C antibody testing for detecting hepatitis C infection among United States veterans?

- The prevalence data, which indicate that 10% of the population has the disease, are needed to determine the positive and negative predictive values of the test.

- A useful method for calculating predictive values for these types of problems is to first create a hypothetical population of any size, 1,000 is usually a good round number, and then to use the prevalence data to first fill in the cells for disease and no disease as shown in table

	Disease		
	Yes	No	Total
Test result			
Positive	*a*	*b*	
Negative	*c*	*d*	
Total	100 (10%)	900 (90%)	1,000 total

Prevalence of disease = $(a + c) / (a + b + c + d)$

The next step is to use the sensitivity and specificity data to fill in cells *a* and *d*.

	Yes	No	Total
Test result			
Positive	Sensitivity = 99% 100 × 0.99 = 99		
Negative		specificity = 95% 900 × 0.95 = 855	
Total	100	900	1,000 total

Now there is enough information to complete the rest of the table

	Yes	No	Total
Test result			
Positive	100 x 0.99 = 99	45	144
Negative	1	900 x 0.95 = 855	856
Total	100	900	1,000 total

Positive predictive value = a / (a + b) = 99 / 144 = 69%

- The negative predictive value of the hepatitis C antibody test is 855/856 × 100% = 99.9%

- The interpretation of this result is, “a U.S. veteran who tests negative for hepatitis C antibody has a 99.9% chance of not having hepatitis C”

- Note that the “true” diagnosis of hepatitis C refers to the use of a gold-standard method.

- The polymerase chain reaction, or PCR test for hepatitis C viral antigen is a gold-standard method that is used for detecting hepatitis C.

Example

- The prevalence of hepatitis C infection among intravenous drug users is 30%. An intravenous drug user undergoes hepatitis C antibody testing and tests positive.

- What is the probability that this person has hepatitis C infection?

Hepatitis C antibody testing: 30% prevalence of disease

	Disease		
	Yes	No	
Test result			
Positive	300 × 0.99 = 297	35	332
Negative	3	700 × 0.95 = 665	668
Total	300	700	1,000 total

Positive predictive value = a / (a +b) = 297 / 332 = 90%

Negative predictive value = d / (c + d) = 665 / 668 = 99.6%

- The positive predictive value of the hepatitis C antibody test among intravenous drug users has now increased to 297/332 = 90%.

- The sensitivity and specificity of the test have remained fixed. Given a positive hepatitis C antibody test result, there is now a 90% chance that this person has hepatitis C.

- This is not surprising since this person had a higher “baseline” risk of hepatitis C prior to antibody testing, due to their use of intravenous drugs.

- The negative predictive value is now slightly lower than that of previous example, again because the “baseline” risk of hepatitis C is higher prior to testing.

EXAMPLE

- A new **ELISA** (antibody test) is developed to diagnose HIV infections. Serum from 80 patients that were positive by Western Blot (the Gold Standard assay) was tested, and 60 were found to be positive by the new ELISA screening test.

EXAMPLE

- The manufacturers then used the new ELISA to test serum from 120 study participants that were negative by Western Blot (the Gold Standard assay); 70 were found to be negative by the new test.

EXAMPLE

- Knowing the prevalence of HIV in country (X) is 22%.
- Calculate:
  - Sensitivity and specificity of the ELISA test
  - PPV and NPV for ELISA used in country (X)

EXAMPLE

		HIV		
		Infected	*Non-infected*	*Total*
ELISA Test	Positive	60 (a = TP)	50 (b = FN)	a + b =110 Total test positive
	Negative	20 (c = FP)	70 (d = TN)	c +d = 90 Total test negative
	Total	80 (a + c) Total infected	120 (b + d) Total not infected	a + b + c + d = 200 Total screened

ACCURACY

ACCURACY

The overall accuracy of a clinical test is the proportion of all tested persons who are correctly identified by the test, that is, the proportion of all test results, both positive and negative, that are correct

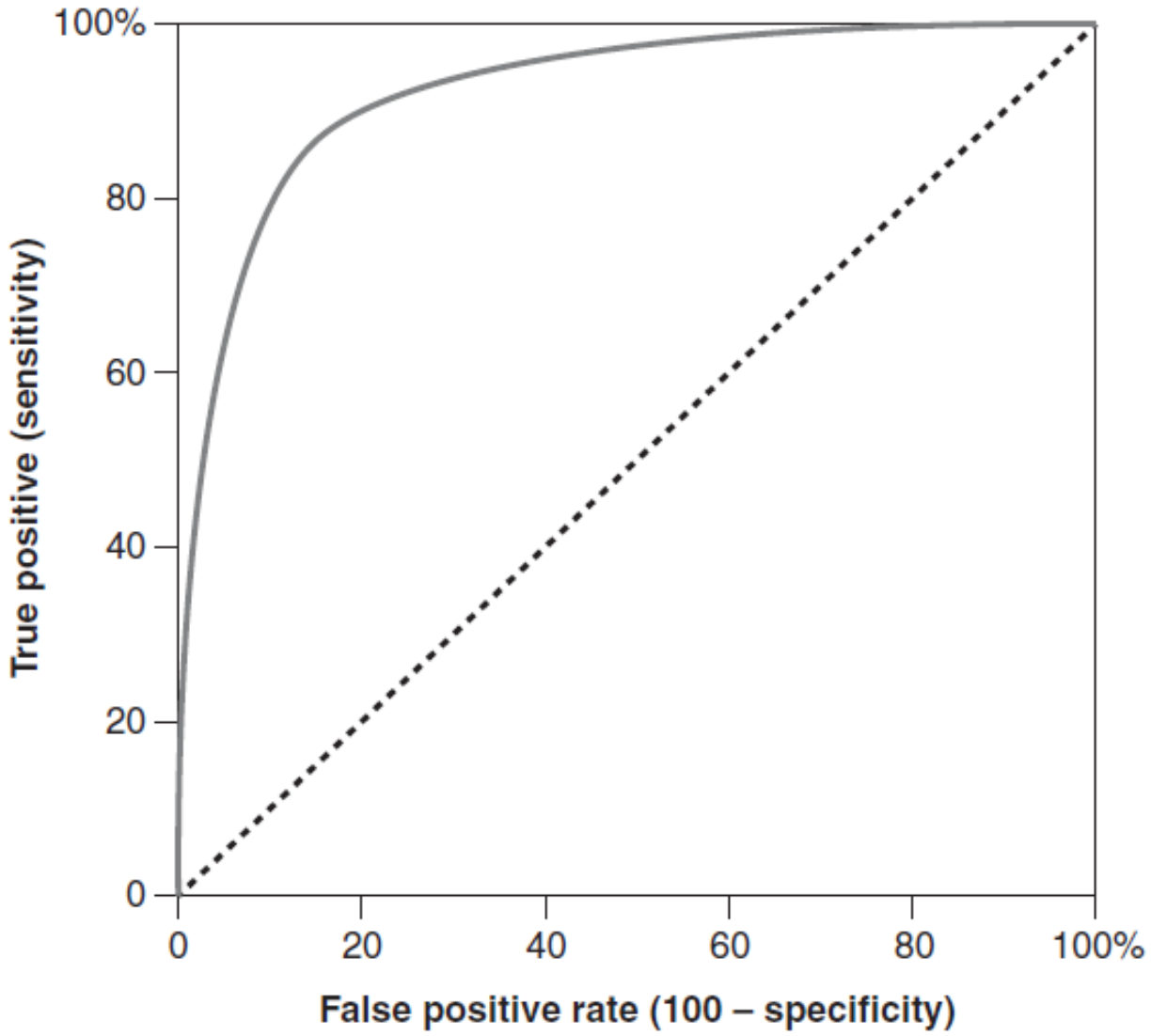
ACCURACY

Accuracy is therefore the number of “true” results (true positives and true negatives) divided by the total of all the test results (true positives, true negatives, false positives, and false negatives)

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

Receiver Operating Characteristic Curve

To help to determine an appropriate cutoff point for a “positive” test, the relationship between sensitivity and specificity can be clarified by plotting a test’s true positive rate (sensitivity) against its false positive rate (100 - specificity) for different cutoff points


100%
80
60
40
20
0
True positive (sensitivity)
0
20
40
60
80
100%
False positive rate (100 – specificity)

Receiver Operating Characteristic Curve

Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular cutoff point or decision threshold. A test that discriminates perfectly between the presence and absence of disease would have an ROC curve that passes through the upper left corner (100% sensitivity, 100% specificity)

Receiver Operating Characteristic Curve

So the closer the curve is to the upper left corner, the higher the overall accuracy of the test. A completely random test (e.g., coin tossing) would give an ROC “curve” that is actually the dashed line

Receiver Operating Characteristic Curve

The shape of the curve therefore reflects the quality of the test; the better the test, the more the curve moves to the upper left. This can be quantified in terms of the area under the curve (AUC); the worst case is 0.5 (the dashed line), and the best is 1.00 (upper left-hand corner)

Receiver Operating Characteristic Curve

A “good” test is one with a high rate of true positives and a low rate of false negatives over a reasonable range of cutoff values; in other words it has a high AUC as the curve moves towards the upper left corner

Receiver Operating Characteristic Curve

As a rule of thumb, an AUC of 0.5 to 0.6 is almost useless, 0.6 to 0.7 is poor, 0.7 to 0.8 is fair, and 0.8 to 0.9 is very good

AUC reflects the test's ability to discriminate between those with and without disease. ROC curves therefore allow different tests and different cutoff points to be compared