

Screening & Diagnostic Tests

**Dr. M.M. Mweu,
MBChB Level II Epidemiology,
21 May, 2018**

Screening

- Defined as *population testing* to identify *early disease* or *precursors* of disease in *asymptomatic* individuals
- Can either be directed at the whole pop (mass screening) or focussed on groups considered to be at high risk of developing the disease (selective screening) e.g. screening to identify individuals with risk factors such as high blood pressure, high cholesterol, smokers etc.
- Distinction has to be made between “**case finding**” – identification of cases with *established disease not* previously *known to health care* – diagnostic tests used to confirm d’s
- Case finding – directed at pop *with d’s* whereas screening is directed at pop *without d’s* (screening tests used)
- It is only worthwhile to screen for d’ses or problems for which there’s an effective intervention i.e. “intention to treat”

Screening

- Examples of screened for d'ses – Breast cancer, cervical cancer

Criteria for appropriateness for screening

- ❑ The condition being screened for should be an important health problem i.e. has high prevalence and results in serious disability/death if untreated
- ❑ Natural history of the disease should be well understood i.e. there's a critical point in the natural history of a d'se before which treatment is likely to be effective e.g. before d'se spread to lymph nodes
- ❑ Treatment at an early stage should be of more benefit than at a later stage
- ❑ There should be a suitable test for the early stage
- ❑ The test should be acceptable e.g. some people may be reluctant to undergo X-rays
- ❑ Intervals for repeating the test should be determined (intervals should ideally be shorter than the *sojourn time* – duration of pre-clinical detectable phase of the d'se)

Screening

- ❑ There should be adequate health service provision for the extra clinical workload resulting from the screen
- ❑ The risks (both physical and psychological) of undergoing screening should be less than the benefits e.g. for invasive procedures
- ❑ The costs should balance against the benefits

Accuracy, precision & agreement of tests

- **Accuracy** refers to ability of test to give a true measure of the substance of interest (e.g blood glucose, serum antibody level) – also is ability of test to detect health/disease
- **Precision** of a test refers to how consistent the results from the test are (variability among test results) – if a test repeatedly gives the same value (regardless of whether/not it's the correct value) it is precise
- Variability among repeated test results from same sample in same lab is repeatability whereas that from testing same sample in different labs is reproducibility
- Results from inaccurate tests can be “corrected” if a measure of the inaccuracy is available and used to adjust test results
- Imprecision is dealt with by performing repeated tests and averaging the results

Accuracy, precision & agreement of tests

Accuracy of tests

- Based on 2 characteristics:
 - **Sensitivity** – ability of a test to detect diseased individuals correctly
 - **Specificity** – ability of a test to detect non-diseased individuals correctly
- In order to assess a new test (determine its Se & Sp) a *gold standard* is required.
- *Gold standard* – test/procedure that is “absolutely” accurate – identifies all d’sed individuals that exist and misdiagnoses none

Accuracy of tests

- In reality, there are very few true gold standards – partly due to imperfections of the test itself but also biological variability e.g. fluctuations in antibody or antigen concentrations in host

	Disease +	Disease -	Total
Test +	a (True positive)	b (False positive)	(a + b)
Test -	c (False negative)	d (True negative)	(c + d)
	(a + c)	(b + d)	a+b+c+d = n

- Sensitivity* (Se) is proportion of diseased individuals that test positive ($\frac{a}{a+c}$) or probability of testing positive given that the individual is diseased $p(T^+|D^+)$
- Specificity* (Sp) is the proportion of non-diseased individuals that test negative ($\frac{d}{b+d}$) or probability of testing negative given that the individual is non-diseased $p(T^-|D^-)$
- $1 - Sp$: false positive fraction; $1 - Se$: false negative fraction

Accuracy of tests

- If you want to confirm d'se (i.e. rule in d'se), a test with *high* Sp is desirable so as to reduce false positives
- If you want to rule out d'se, a test with a *high* Se is desirable so as to reduce false negatives
- Se and Sp are population-specific and may vary across pops

True and Apparent prevalences

- **True prevalence** (Tp) is the proportion of the population that is truly diseased ($\frac{a+c}{n}$). Tp also called **pre-test prevalence**
- However, unless test is perfect, test results only provide estimate of Tp – **apparent prevalence** (Ap) – is the proportion of tested individuals that are positive ($\frac{a+b}{n}$). Ap also called **post-test prevalence**. If Tp , Se and Sp are known:

$$Ap = Se * Tp + (1 - Sp) * (1 - Tp)$$

Predictive values

- Se & Sp may not be useful when applied to individuals of unknown disease status
- Often times we want to know the probability that an individual has/does not have d'se in question given that they test positive/negative – these probabilities are **predictive values** – affected by Tp, Se and Sp of the test
- **Positive predictive value (PV+)** – probability that an individual has the d'se given that he/she tests positive
 $p(D^+|T^+)$ or $(\frac{a}{a+b})$
- **Negative predictive value (PV-)** – probability that an individual doesn't have the d'se given that he/she tests negative
 $p(D^-|T^-)$ or $(\frac{d}{c+d})$

Predictive values

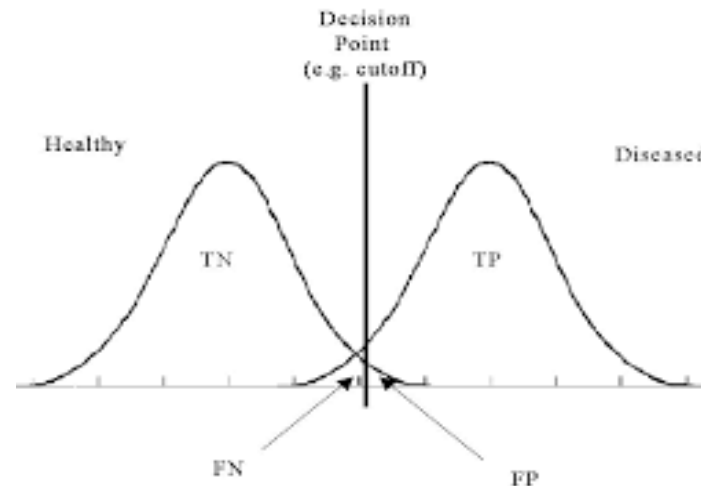
- PPV in a population can be increased by:
 - Test individuals in a pop where the prevalence is high e.g. in a screening programme, testing individuals that are likely to have the d'se – those with a particular history such persistent coughing in TB
 - Using a test with a high Sp
 - Using more than one test with an individual being considered diseased if positive in all tests

Interpreting test results on a continuous scale

- Sometimes substance being measured e.g. antibody titres, serum calcium, liver enzymes, is recorded on a continuous scale
- Hence need to select a cut-point (cut-off/threshold) to determine level of result that indicates positive test result
- Often though there's an overlap in the distribution of the substance being measured between healthy & d'sed individuals

Interpreting test results on a continuous scale

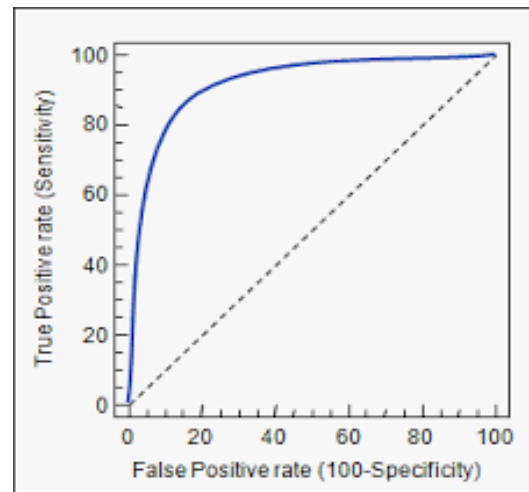
- A cut-point that optimises Se and Sp of the test is selected but with dilemma as shown:



- If there's an overlap in test values for D+ & D- whatever cut-point is selected will result in both FP and FN test results
- If the cut-off is raised, Sp increases (FPs decrease) and Se decreases (more FNs). Lowering cut-off has opposite effect
- To choose among multiple cut-points, graphical procedures e.g. the **receiver operating characteristic curves (ROC)** are used to determine optimal cut-point

Receiver operating characteristic (ROC) curves

- ROC curve is a plot of the true positive rate (Se) of test against the false positive rate ($1 - Sp$) computed at a no. of different cut-points to select the optimal cut-point for distinguishing between d'sed and non-d'sed individuals



- The 45° line represents test with discriminating ability no better than chance
- The closer the ROC curve gets to top-left corner of graph, the better the ability to differentiate between d'sed & non-d'sed

Receiver operating characteristic (ROC) curves

- Top-left corner represents test with Se and Sp of 100%
- Assuming equal costs of FN and FP test results, the optimum cut-point is that with $Se + Sp$ at a maximum where curve gets closest to top left corner of the graph (or furthest away from 45° line)

Example of Se , Sp , Ap , TP , $PV+$, $PV-$ calculation

	D+	D-	Total
T+	178 (a)	622 (b)	800 (a+b)
T-	55 (c)	818 (d)	873 (c+d)
Total	233 (a+c)	1440 (b+d)	1673 (n)

$$Se = 178/233 = 76.4\%$$

$$Sp = 818/1440 = 56.8\%$$

$$Tp = 233/1673 = 13.9\%$$

$$Ap = 800/1673 = 47.8\%$$

$$PV+ = 178/800 = 22.3\%$$

$$PV- = 818/873 = 93.7\%$$