Evidence Based Medicine:
Module 9 - Diagnostic and Screening Studies

Key Points:

Screening Test:

Bring forward in time the diagnosis (prior to the onset of symptoms)
Case broad net to identify those at high risk for disease
Often applied to a broader population with a lower a priori probability of disease

Minimize false negatives *even if there are false positives*. Thus, screening tests are NOT used to diagnose any condition. Any positive screening test **MUST** be followed by a more SPECIFIC test in order to make a diagnosis.

Diagnostic Process

- Method of exhaustion (medical student reasoning)
- Pattern recognition (experienced clinician)
- Hypothetical-deductive reasoning
    1. Estimate a pre-test probability (likelihood of a specific diagnosis prior to application of a diagnostic test
    2. Decide whether a diagnostic test is required. Has the 'treatment threshold' been crossed? Has the 'test threshold' been crossed?
    3. Apply the test
    4. Estimate a post-test probability

"Treatment threshold"

Is the provisional diagnosis so likely that we can move on to treatment/management?
If yes, further testing is NOT necessary

"Test threshold"

Is a specific diagnosis deemed so unlikely that we can comfortably dismiss it?
If no, then further testing is necessary

Post-test Probability

Start with a reasonable estimate of pre-test probability
Apply an accurate diagnostic test
Use the combined information from the pre-test probability and the accuracy of the diagnostic test, to estimate the post-test probability

Index Test:

The new test that will be used in clinical practice to differentiate the condition of interest from some other state

Reference Standard

Gold standard or the 'truth'
The best available procedure, method or criteria used to establish the presence or absence of the condition of interest
Should be independent of the index test

Diagnostic Test Metrics

- Sensitivity and Specificity
- Positive & Negative Predictive Values
- Positive & Negative Likelihood Ratios
- Precision
- Confidence Intervals
- Reproducibility

Sensitivity: among people who have the disease, the proportion who test positive.

Sensitivity is defined as (true positive) / (true positive + false negative)

Specificity – among people who do not have the disease, the proportion who test negative

Specificity is defined as (true negative) / (true negative + false positive)

Receiver Operator Characteristic (ROC) Curves

Plot of sensitivity vs (1-specificity)
Demonstrates the Trade off between sensitivity and specificity, and different "cut off points).
Demonstrates the Trade off" between true positives and false positives

Baye's Theorem and the Predictive Value of a Positive Test

The probability of a test demonstrating a true positive depends not only on the sensitivity and specificity of a test, but also on the prevalence of the disease in the population being studied.

The chance of a positive test being a true positive is markedly higher in a population with a high prevalence of disease.

In contrast, if a very sensitive and specific test is applied to a population with a very low prevalence of disease, most positive tests will actually be false positives

Using Positive Predictive Values derived in one setting will likely not be valid in another

setting with different disease prevalence. This can be better addressed using Likelihood Ratios

Likelihood Ratios

The likelihood ratio (LR) is the most clinically useful measure of diagnostic test accuracy. The LR associated with a test indicates by how much a given test result will raise or lower the pretest probability of the target disorder

A LR of 1 means that the post-test probability is the same as the pretest probability

LR values >1.0 increase the probability that the target disorder is present, and the higher the likelihood ratio, the greater is this increase.

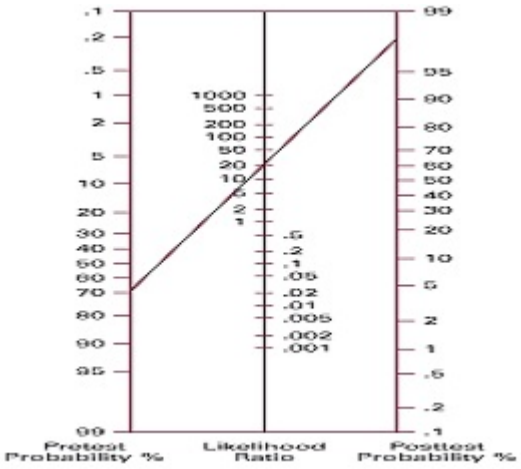
Conversely, LR values <1.0 decrease the probability of the target disorder, and the smaller the likelihood ratio, the greater is the decrease in probability and the smaller is its final value.

Advantages of LR include that they can used directly in clinical reasoning to establish post-test probability and they are not affected by disease prevalence.

$$LR = \frac{\text{Likelihood of a test result when disease is present}}{\text{Likelihood of a test result when disease is absent}}$$

$$\text{LR positive} = \frac{\text{sensitivity}}{\text{1-specificity}} \qquad \text{LR negative} = \frac{\text{1-sensitivity}}{\text{specificity}}$$

Likelihood Ratio	Change in Probability	Clinical Importance
> 10 or < 0.1	Large	Often very high
5-10 or 0.1-0.2	Moderate	Moderately high
2-5 or 0.2-0.5	Small	Sometimes
1-2 or 0.5-1.0	Very small	Rare



Fagan's Normogram: estimating a pre-test probability and then applying a test (knowing its likelihood ratio) results in the ability to estimate the post-test probability.

Agreement

Many tests require observer interpretation. Clinical utility and generalizability are affected by the inter-observer agreement. If there is poor agreement about a test result between different observers or in different situations, the test may not be useful.

Agreement is measured with the **kappa (κ) statistic**

Kappa expresses the extent of the possible agreement over and above chance
If the raters agree on every judgment, the total possible agreement is always 100%
Expressed in a range from -1.0 (perfect disagreement) to 1.0 (perfect agreement)
Here are some useful numbers to aid in interpretation of the kappa statistic

< 0 - no agreement
0.0 – 0.20 - slight agreement
0.21 – 0.40 - fair agreement
0.41 – 0.60 - moderate agreement
0.61 – 0.80 - substantial agreement
0.81 – 1.00 - almost perfect agreement