Critical Appraisal Course for Emergency Medicine Trainees

Module 5

Evaluation of a Diagnostic Test

Evaluating a diagnostic test

- The reference standard
- Independence of the reference standard
- Blinding
- Patient spectrum
- Inter-observer error
- Sensitivity, specificity, PPV, NPV
- Likelihood ratios

The reference (gold) standard

- Used to determine whether the patient has the disease or not
- Possible reference standards:
  - A single accurate diagnostic test
  - A combination of tests
  - Tests with follow-up for negatives
- Often involves trade-off between accuracy and feasibility of reference standard

Independent reference standard

- Ideally all patients should get the same reference standard test
- If not, the results of the test under investigation should not determine which reference standard test is used
- Work-up bias: Patients with +ve test get one reference standard, those with –ve test get another
- Incorporation bias: The test under investigation is incorporated into the reference standard test

Blinding

- Was the test under investigation interpreted blind to the results of the reference standard?
- Was the reference standard interpreted blind to the results of the test under investigation?
- Awareness either way may influence interpretation

Patient spectrum

- Study population should be representative of the population in which the test will be used
- How were patients selected?
- What was disease prevalence?
- High prevalence suggests a selected population

Inter-observer error (reliability)

- Does the test give the same results when interpreted by different people?
- Measurement of reliability needs to take into account agreement due to chance
- Usually measured by Kappa
- Kappa=0: any agreement was due to chance
- Kappa=1: perfect agreement

What do the following mean?

- Case positive
- Case negative
- Test positive
- Test negative
- Prevalence

What do the following mean?

- Case positive: person with disease (positive reference standard)
- Case negative: person without disease (negative reference standard)
- Test positive: person with positive test under investigation
- Test negative: person with negative test under investigation
- Prevalence: proportion with disease

What do the following mean?

- True positive
- True negative
- False positive
- False negative

What do the following mean?

- True positive: positive test result & has disease
- True negative: negative test result & does not have disease
- False positive: positive test result but does not have disease
- False negative: negative test result but has disease

What do the following mean?

- Sensitivity
- Specificity
- Positive predictive value
- Negative predictive value

What do the following mean?

- Sensitivity: proportion of patients with disease who have a positive test
- Specificity: proportion of patients without disease who have a negative test
- Positive predictive value: proportion of patients with a positive test who have disease
- Negative predictive value: proportion of patients with a negative test who do not have disease

How are parameters used?

- Sensitivity: If a sensitive test is negative disease is ruled out (SnOut)
- Specificity: If a specific test is positive disease is ruled in (SpIn)
- These are useful, simple rules but not strictly true
- Likelihood ratios are better for decision-making

Prevalence

- Sensitivity and specificity are mathematically independent of prevalence
- Positive predictive value increases when prevalence increases
- Negative predictive value decreases when prevalence increases

Likelihood ratio

- Applies to a piece of diagnostic information, such as an observation, a clinical finding or a test result
- Tells you how diagnostically useful that piece of information is
- Is a number between zero and infinity
- If greater than one, indicates that the information increases the likelihood of the suspected diagnosis
- If less than one, indicates that the information decreases the likelihood of the suspected diagnosis

Diagnostic value of likelihood ratios

1	None at all
0.5 to 2	Little clinical significance
2 to 5	Moderately increases likelihood of disease, but does not rule-in
0.2 to 0.5	Moderately decreases likelihood of disease, but does not rule-out
5 to 10	Markedly increases likelihood of disease, may rule-in
0.1 to 0.2	Markedly decreases likelihood of disease, may rule-out
>10	Rules in
<0.1	Rules out

Examples: Clinical assessment for DVT

Feature	Likelihood ratio of positive finding	Likelihood ratio of negative finding
Past history VTE	2.5	0.88
Malignancy	2.6	0.88
Immobilisation	1.9	0.89
Recent surgery	1.7	0.93
Difference in calf diameter	1.8	0.51
Homan's sign	1.4	0.87
Oedema	1.2	0.89

Calculating likelihood ratios

- LR of +ve test = sensitivity / (1-specificity)
- LR of –ve test = (1-sensitivity) / specificity

Summary

- What reference standard was used?
- Was it independent of the test under investigation?
- Was the test under investigation interpreted blind?
- Was the reference standard interpreted blind?
- Was an appropriate spectrum of patients included?
- Was inter-observer error measured?
- What were the sensitivity & specificity?
- What were the likelihood ratios?

Any questions or comments?