Basic concepts of statistical analysis

Sasivimol Rattanasiri, Ph.D

Section for Clinical Epidemiology and Biostatistics

Ramathibodi Hospital, Mahidol University

E-mail: sasivimol.rat@mahidol.ac.th

Outline of talk

Diagnostic tests:

- Categorical data
  - Estimate sensitivity, specificity
  - Estimate LR^+ , LR^-
- Continuous data
  - Estimate area under receiver operating characteristics (ROC) curve

Diagnostic study

- Study design: Cross-sectional or case-control study
- Study population should be identified with some demographics and diagnostic characteristics
- The criterion (gold) standard represents the truth, or as close to the truth as current measurements
- Types of data of diagnostic tests: continuous or categorical data

Example

Hysteroscopy was used to diagnose uterine cancer in premenopausal women. The gold standard to diagnose uterine cancer was pathology analysis.

- Study design: cross-sectional study
- Study population: premenopausal women
- Gold standard: Pathology
- Diagnostic test: Hysteroscopy

Characteristics of diagnostic test

Test result	Actual condition of population: With disease	Without disease
Positive	a (True positive/ (Sensitivity)	b (False positive)
Negative	c (False negative/ specificity)	d (True negative)
Total	a+c	b+d

Sensitivity/Specificity

- **Sensitivity is proportion of true positive**
  - sensitivity=a/(a+c)
- **Specificity is proportion of true negative**
  - specificity=d/(b+d)

PPV/NPV

- Positive predictive value (PPV) is number of diseased patients with positive tests divided by number of patients with positive tests
  - PPV=a/(a+b)
- Negative predictive value (NPV) is number of non-diseased patients with negative tests divided by number of patients with negative tests
  - NPV=d/(c+d)

		Patients with bowel cancer (as confirmed on endoscopy)		
		Condition positive	Condition negative	
Fecal occult blood screen test outcome	Test outcome positive	**True positive** (TP) = 20	**False positive** (FP) = 180	Positive predictive value = TP / (TP + FP) = 20 / (20 + 180) = **10%**
	Test outcome negative	**False negative** (FN) = 10	**True negative** (TN) = 1820	Negative predictive value = TN / (FN + TN) = 1820 / (10 + 1820) ≈ **99.5%**
		Sensitivity = TP / (TP + FN) = 20 / (20 + 10) ≈ **67%**	Specificity = TN / (FP + TN) = 1820 / (180 + 1820) = **91%**	

Likelihood ratio positive

$$LR^{+} = \frac{\text{likelihood of positive result in patients with disease}}{\text{likelihood of positive result in patients without disease}}$$

- LR^{+} =sensitivity/(1-specificity)
- LR^{+} =6.2 means that a positive test results in 6.2 times more likely to occur in patients with disease than in patient without disease.

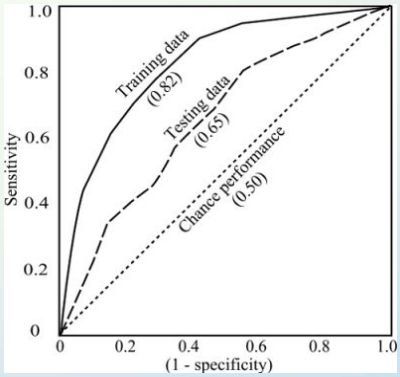
Likelihood ratio negative

$$LR^{+} = \frac{\text{likelihood of negative result in patients with disease}}{\text{likelihood of negative result in patients without disease}}$$

- LR^{-} = (1-sensitivity)/specificity
- LR^{-} =0.45 means that a negative result is 55% less likely to occur in a patient with disease than in a patient without disease

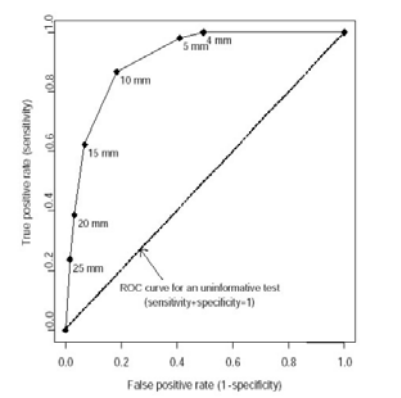
ROC curve

• The receiver operating characteristics (ROC) curve is a graph in which the Y axis represents sensitivity and the X axis represents 1-specificity.



ROC curve

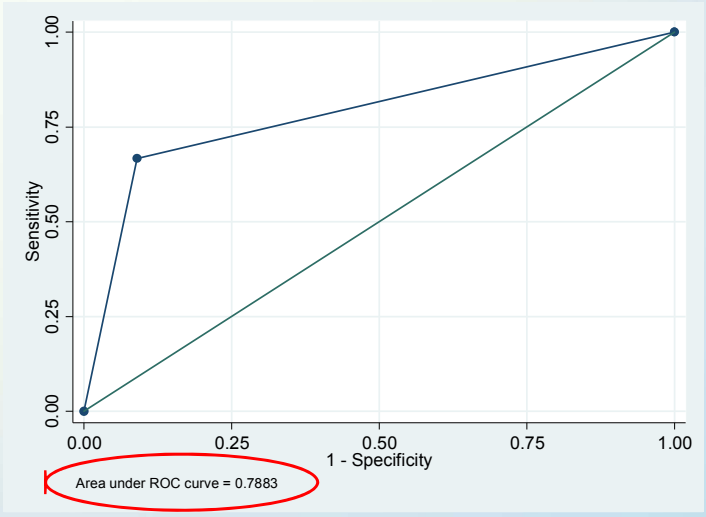
The area under ROC curve indicates the accuracy of the test with different cut-points.



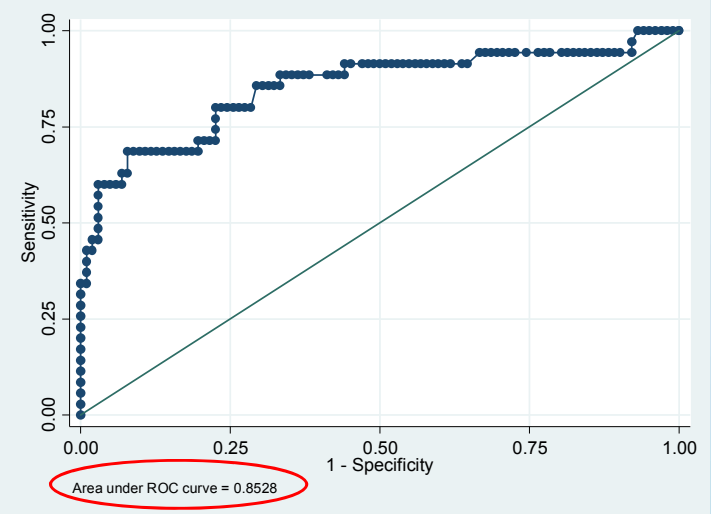
Diagnostic test of bowel cancer

```
. diagt endoscopy fob_test

           |       fob_test
 endoscopy |      Pos.       Neg. |     Total
-----------+----------------------+----------
  Abnormal |        20         10 |        30
    Normal |       180      1,820 |     2,000
-----------+----------------------+----------
     Total |       200      1,830 |     2,030

True abnormal diagnosis defined as endoscopy = 1


                                                        [95% Confidence Interval]
-------------------------------------------------------------------------------
Prevalence                          Pr(A)       1.5%        1%        2.1%
-------------------------------------------------------------------------------
Sensitivity                        Pr(+|A)      66.7%     47.2%      82.7%
Specificity                        Pr(-|N)        91%     89.7%      92.2%
ROC area                 (Sens. + Spec.)/2       .788      .702       .874
-------------------------------------------------------------------------------
Likelihood ratio (+)     Pr(+|A)/Pr(+|N)       7.41      5.55       9.89
Likelihood ratio (-)     Pr(-|A)/Pr(-|N)       .366      .221       .608
Odds ratio                     LR(+)/LR(-)       20.2      9.47       43.1
Positive predictive value        Pr(A|+)        10%     6.22%        15%
Negative predictive value        Pr(N|-)      99.5%       99%      99.7%
-------------------------------------------------------------------------------
```


ROC curve for diagnostic test of bowel cancer



ROC curve for PSA level to diagnose prostate cancer



Outline of talk

Test for agreement:

- Categorical data
  - Kappa analysis
- Continuous data
  - Level of agreement

Inter-rater agreement

- Agreement between *categorical* assessments is usually considered as comparing the ability of different raters to classify subjects into one of several groups.
- For example, we would like to assess the classification by two radiologists of 85 xeromammograms as
  - Normal
  - Benign
  - Suspicion of cancer
  - Cancer

Strength of agreement

Value of kappa	Strength of agreement
< 0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1.00	Very good

Level of agreement

- In clinical measurement comparison of a new measurement technique with an established one is often needed to see whether they agree sufficiently for the new to replace the old.
- For example, the PEER values were measured by two different methods; large and mini peak flow meters. If the two PEER meters were differed by more than 10 l/min, we could replace large meter by mini meter because small difference would not affect decisions on patient management.