THE MAIN MEASURES OF ASSOCIATION IN EPIDEMIOLOGY:

		DISEASE		
EXPOSURE		**YES**	**NO**	Marginal total
	YES	**A** Exposed, and diseased	**B** Exposed, Not diseased	**A+B** Exposed
	NO	**C** Not exposed, diseased	**D** Not exposed, Not diseased	**C+D** Non-exposed
	Marginal total	**A+C** Diseased	**B+D** Non-diseased	Grand total (n) = **A+B+C+D**

2X2 TABLE (CONTINGENCY TABLE): The main tool that we are going to use

Measures of Association (to look for an association between a certain risk factor & a disease)

Chi-square: in Cross-sectional studies

- Tests whether there is an association between two categorical variables like dichotomous variable
- we usually use a software like SPSS or excel to calculate chi-square just by providing the cells
- The important thing is the P-value (statistical significance)
- If p-value<0.05, there is a significant association between the risk factor and the disease
- Chi-square statistic tells only whether there is association or no association. It doesn't tell us how strong an association is

Relative risk (RR) or Risk Ratio (RR): In a cohort study

RR= Incidence (risk) among exposed
Incidence (risk) among non-exposed
= A/(A+B)
D/(D+C)

- The estimation of disease risk associated with exposure (indication of strength of association)
- Can measure incidence
- RR = 10/1=10 → interpretation: The exposed have 10 times the risk of developing the disease when compared to non-exposed
- *High* association if RR>3 *Moderate* if RR is between 1.5 & 2.9 *Weak* association if RR is between 1.2 & 1.4 *No association* exists if RR is 1 *Negative* association (protective effect) if RR <1

ODDS RATIO (OR): case-control studies

case control studies does not provide incidents→ cannot use RR BUT OR can be a good estimate of RR

OR= Odds of exposure among diseased
Odds of exposure among non-diseased

OR=(a/c) = ad/bc (cross product of the table)
(b/d)

Same interpretation as RR: *High(strong)* association if OR>3 *No association* exists if OR is 1 *Negative* association (protective effect) if OR <1

RR can be best estimated by OR if the following conditions are fulfilled: 1. Controls are representative of general population 2. Selected cases are representative of all cases in the population 3. The disease is rare (the more rare the disease is the more close would the OR be to the RR)

Attributable Risk (AR): (risk difference)

AR= Risk (incidence) in exposed (-) Risk (incidence) in non-exposed

$$AR = \frac{a}{(a+b)} - \frac{c}{(c+d)}$$

Attributable risk percent (AR%) the proportion of disease among the exposed that is attributable to the exposure.

AR%= (Risk in exposed – Risk in non-exposed)X100%
Risk in exposed

AR% suggests the amount of disease that might be eliminated if the exposure could be controlled or eliminated

Interpretation of the numbers:

AR=0 → No association, AR>0 → Positive association, AR< 0 → Negative association

A **risk factor** is any factor positively associated with a disease (RR>1).

A preventive factor is any factor negatively associated with a disease (RR<1).

Risk and preventive factors may or may not be amenable to change; we cannot change the person genetic makeup or change a person's age, but we can change eating habits or smoking

Population Screening:

An application of a test to asymptomatic people to detect occult disease [subclinical disease] or a precursor state.

Immediate objective of a screening test – to classify people as being likely or unlikely of having the disease.
Ultimate objective: to reduce mortality and morbidity → To find diseased people in early stages

Many countries have national screening programs for various diseases: cancer screening, hypertension and diabetes, newborns for phenylketonuria (PKU) to prevent mental retardation

Cancer screening examples: Pap smears to detect cervical neoplasia, mammography and physical breast exam to detect breast cancer, PSA to detect early prostate cancer, and fecal occult blood testing to detect colorectal cancer or adenomas

There are two major dimensions to accuracy:

1. Reliability (consistent): get same result if we repeat the test again (However, a test can be reliable but still give an incorrect result). Reliability does not ensure validity, but lack of reliability constrains and jeopardizes validity.

2. Validity: give the correct result (reflect the true state). There are two desirable properties for a screening test related to validity:

1. **Sensitivity of the test**: the probability of correctly classifying someone who has the disease (case).

   Cases found by the test as(positive) + / all cases

2. **Specificity of the test**: the probability of correctly classifying someone without the disease (healthy, non-case).

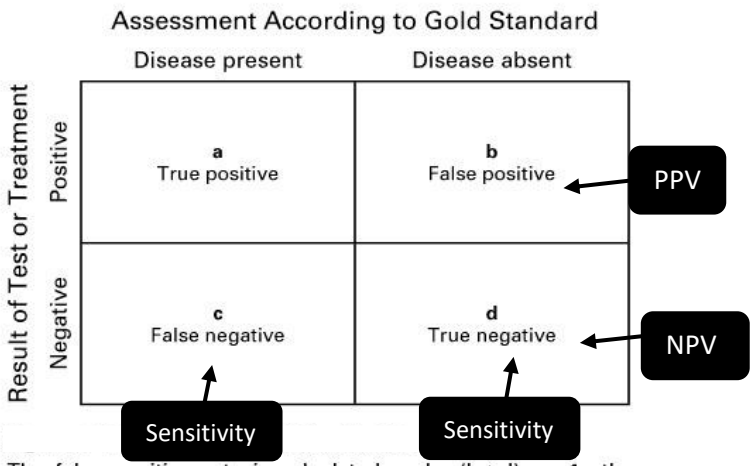
   the number of (non-cases/negative) - / all non-cases

→ Sensitivity and specificity are both probabilities of correct classification of cases and non-cases by screening test.

Probability (proportion) of those tested who are correctly classified by the test:

Positive predictive value = Cases identified / all positive tests

Negative predictive value = Non-cases identified / all negative tests



The false positive rate is calculated as $b \div (b+d)$, or 1 – the specificity; the true positive rate (sensitivity) as $a \div (a+c)$; the true negative rate (specificity) as $d \div (b+d)$, or 1 – the false positive rate; and the positive predictive value as $a \div (a+b)$.

Done by Ghazal Al-Attiyat