Biostats Review: PCOM
March 5, 2016
Parul Chaudhri, DO

Measures that define the **association** between exposures and outcome

- Depends upon the type of study design

1. Relative Risk : the ratio of the incidence of disease in the exposed group divided by the corresponding incidence of disease in the unexposed group

- **Study Design**:
  - Outcome occurs in the future

2. Odds Ratio : odds of exposure in the group with disease divided by the odds of exposure in the control group

- **Study Design**:
  - Outcome occurred in the past

		Disease	
		+	-
Test	+	A	B
	-	C	D

OR= A/ B = AD **RR= A/ (A+ B)**
C/D BC **C/ (C+D)**

Study Design

1. **Cohort Study** :Observational study in which subjects with an exposure of interest (e.g. hypertension) and subjects without the exposure are identified and then followed forward in time to determine outcomes (e.g. stroke).
2. **Case-Control Study**: Observational study that first identifies a group of subjects with a certain disease and a control group without the disease, and then looks to back in time (e.g. chart review) to find exposure to risk factors for the disease.
   - This type of study is well suited for rare diseases.
3. **Cross-Sectional Study** Observational study that is done to examine presence or absence of a disease or presence or absence of an exposure at a particular time. Since exposure and outcome are ascertained at the same time, it is often unclear if the exposure preceded the outcome.
4. **Case Report or Case Series**: Descriptive study that reports on a single or a series of patients with a certain disease. This type of study usually generates a hypothesis but cannot test a hypothesis

Measures of Diagnostic Accuracy

		Disease				
		+	-			
Test	+	True Positive	False Positive	All with Positive Test TP+FP	Positive Predictive Value=	Post-Test Probability Given Positive Test =**PPV**
	-	False Negative	True Negative	All with Negative Test FN+TN	Negative Predictive Value=	Post-Test Probability Given Negative Test =100%-NPV
		All with Disease TP+FN	All without Disease FP+TN	Everyone=**N** TP+FP+FN+TN		
		Sensitivity= TP / (TP+ FN)	**Specificity=** TN / (TN+ FP)	**Pre-Test Probability** (TP+FN)/(TP+FP+FN+TN)		

http://www.musc.edu/dc/icrebm/2x2table.html

Sensitivity (Sn): ability of the test to identify correctly those who have the disease or the percentage of people with the disease who test positive

- High sensitivity helps rule OUT disease (SNOUT)

Specificity (Sp): ability of the test to identify correctly those who do not have the disease or percentage of people without disease who test negative

- High specificity helps rule IN disease (SPIN)

Sn= TP / (TP+ FN) **Sp= TN / (TN + FP)**

Positive Predictive Value: Probability of disease in a patient with a positive test.

Negative Predictive Value: Probability that the patient does not have disease if he has a negative test result.

PPV: TP / (TP + FP) **NPV: TN / (TN + FN)**

Likelihood Ratio: How much a given diagnostic test result will raise or lower the odds of having a disease relative to the prior probability of disease

- independent of disease prevalence

Positive Likelihood Ratio: Odds of disease if a test result is positive
Negative Likelihood Ratio: Odds of disease if a test result is negative

$$\textbf{+ LR} = \frac{\textbf{Sn}}{\textbf{(1- Sp)}}$$

$$\textbf{- LR} = \frac{\textbf{(1- Sn)}}{\textbf{Sp}}$$

Making Inferences About Data

Type I error (alpha): the probability of incorrectly concluding there is a statistically significant difference in the population when none exists.

Type II error (beta): the probability of incorrectly concluding that there is no statistically significant difference in a population when one exists.

Power: is a measure of the ability of a study to detect a true difference.

$$\textbf{Power} = \textbf{1- } \beta$$

95 % Confidence Interval:

- The confidence interval is a measure of variance and is derived from test data.

P-value is a level of statistical significance, and characterized the likelihood of achieving the observed results of a study by chance alone.

- The p-value in and of itself says nothing about the truth or falsity of the null hypothesis.

Prevalence is the existence of disease in the current population.

$$\text{Prevalence} = \frac{TP + FN}{N}$$

Incidence is the occurrence of new cases of disease in a population over a defined time period.

Relative risk is the risk of an event in the experimental group versus the control group in a clinical trial.

Attributable Risk: excess risk of disease in those exposed taking into account the background rate of disease

Absolute Risk Reduction (ARR): absolute adverse event rate for placebo minus the absolute adverse event rate for treated patients

Number needed to treat (NNT): The number of patients who would need to be treated to prevent one adverse outcome is often used to present the results of randomized trials

- The number needed to treat is useful for evaluating data regarding treatments, not diagnosis

NNT= 1/ ARR

Bias: Any systematic error in the design or conduct of a study that results in a mistaken estimate of an exposure's effect on risk of disease.

1. **Lead-time bias** is when a screening test identifies a cancer earlier, thereby increasing the time between diagnosis and death without actually prolonging life.

2. **Length-time bias** is when a screening test finds a disproportionate number of cases of slowly progressive disease and misses the aggressive cases, thereby leading to an overestimate of the effectiveness of the screening.

References:

1. Fletcher RW, Fletcher SW: *Epidemiology: The Essentials*, ed 4. Lippincott Williams & Wilkins, 2005
2. http://www.musc.edu/dc/icrebm/2x2table.html
3. http://www.medpagetoday.com/lib/content/Medpage-Guide-to-Biostatistics.pdf