RESEARCH METHODOLOGY LECTURE-I

Dr. Rizwana Shahid
Assistant Professor Community Medicine
Rawalpindi Medical University, Rawalpindi

Learning Objectives

By the end of lecture, students should be able to:

1. Explain with example concept of correlation and association in research data
2. Distinguish clinical significance from statistical significance
3. Explain principles of correlation analysis for comparing two Continuous variables in same subjects in given data set.
4. Compute co-efficient of correlation and interpret results
5. Compute and interpret determination of correlation for a given data set.
6. Draw & interpret scatter diagrams with respect to the types of correlation

Measures of Association

- A **measure of association** quantifies the relationship between exposure and disease among the two groups. ... Examples of **measures of association** include
  - Odd's ratio (OR)
  - Relative Risk (RR)
  - Attributable Risk (AR)

Odds Ratio in Case-Control Studies (cont.)

BLE 10-9. *Example of Calculating an Odds Ratio From a Case-Control Study*

		First Select	
		CHD Cases	Controls
Then Measure Past Exposure	Smokers	112 (*a*)	176 (*b*)
	Nonsmokers	88 (*c*)	224 (*d*)
	Totals	200 (*a* + *c*)	400 (*b* + *d*)
	Proportions smoking cigarettes	56%	44%

$$\text{Odds ratio } \frac{ad}{bc} = \frac{112 \times 224}{176 \times 88} = 1.62$$

Contingency table (2×2 table)

		Outcome: Yes	Outcome: No
Exposure	Yes	A	B
	No	C	D

Risks Relative Risks and Attributable Risks

Factors	Incidence		Total
	Yes	No	
Exposed group	A	B	A+B
Non-exposed group	C	D	C+D

How many times factor exposure would increase the incidence of an individual:

$$\text{Relative risk} = \frac{\text{Incidence risk among an exposed group}}{\text{Incidence risk among a non-exposed group}} = \frac{\frac{A}{A+B}}{\frac{C}{C+D}}$$

Relative risk larger than 1 represents that risks have increased due to factor exposure.

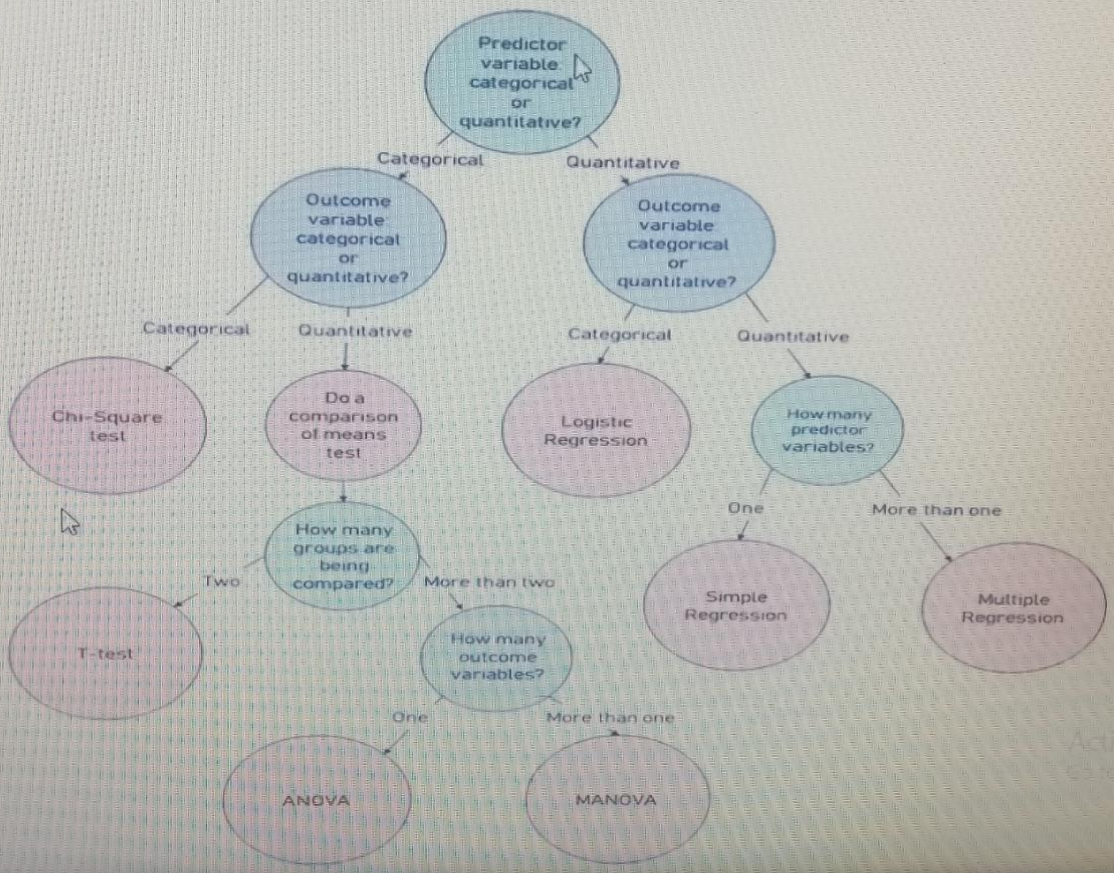
The value obtained by subtracting 1 from the relative risk is an excess relative risk, showing an increased amount of risks.

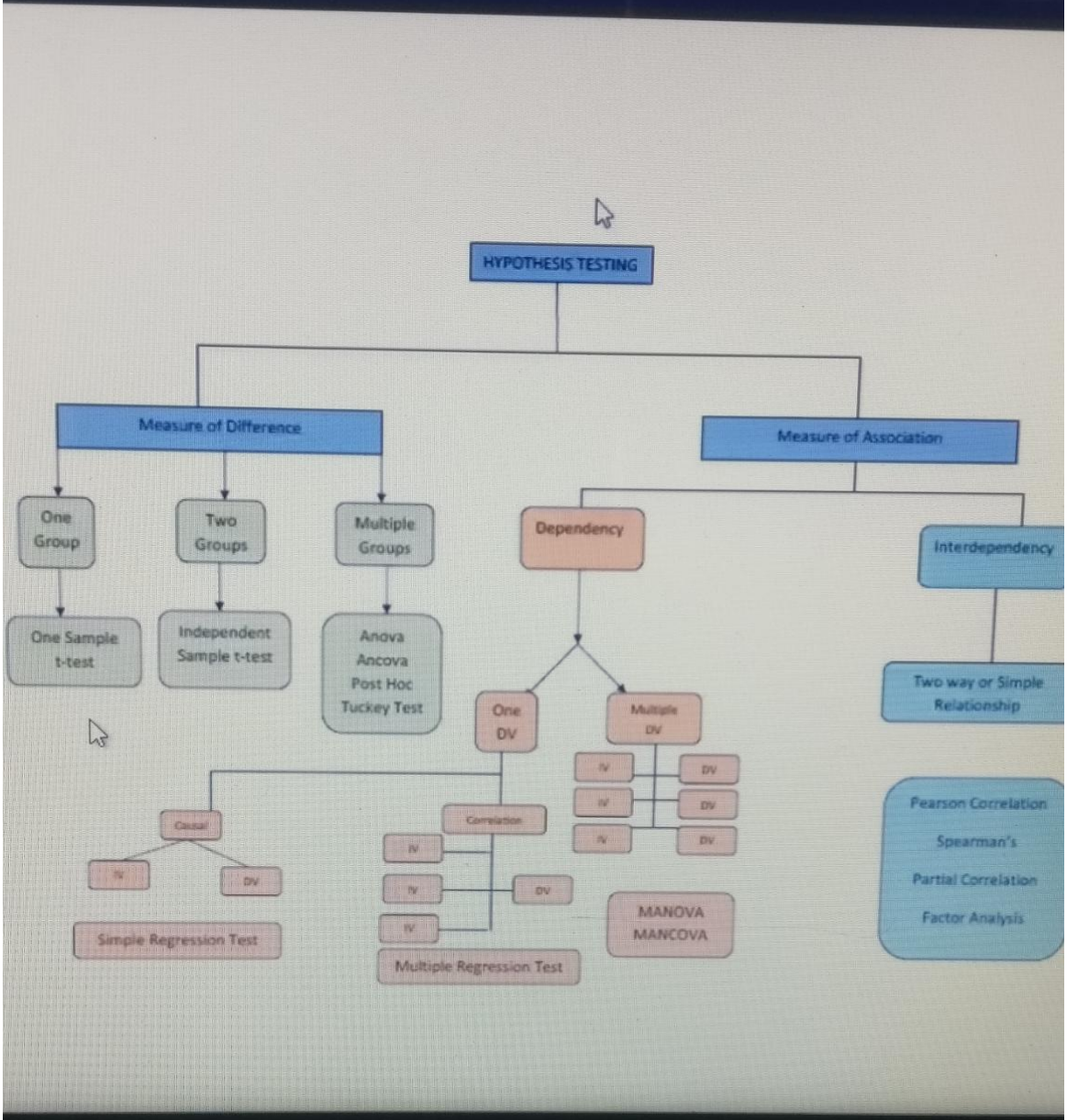
How many times factor exposure would increase the incidence rate of a group:

$$\text{Attributable risk} = \text{Incidence risk among an exposed group} - \text{Incidence risk among a non-exposed group}$$

$$= \frac{A}{\quad} - \frac{C}{\quad}$$

Choosing a statistical test





Correlation & Association

- **Correlation:**

It is the relationship between dependent and independent variables.

- **Association:**

A vague term used to describe relationship between 2 quantitative variables.

❖ Correlation shows linear relationship where as association does not need to be linear.

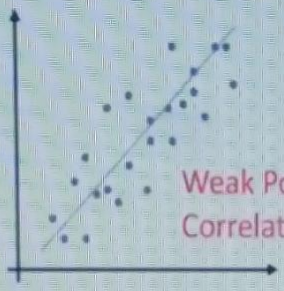
Correlation

- The measure of correlation is called correlation co-efficient.
- The degree of correlation is expressed by co-efficient which is from -1 to +1.
- Direction of correlation is indicated by the sign.

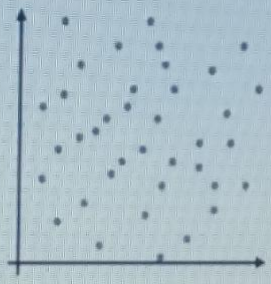
Types of Correlation:

- Positive correlation
- Negative correlation
- No correlation

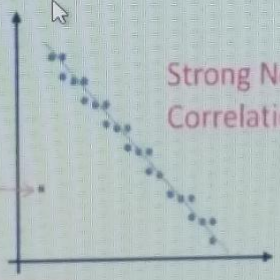
Scatter Diagram



Weak Positive Correlation

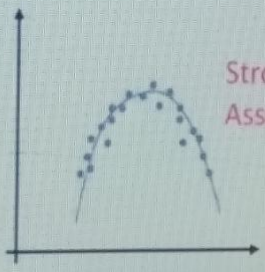


No Correlation



Strong Negative Correlation

This point seems to be an outlier.



Strong Association

No Correlation, because the relationship is not linear.

- Correlation co-efficient (r)

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

- Co-efficient of Determination (r^2)

The amount of variance in dependent variable that is predictable from variance in independent variable

Types of correlation

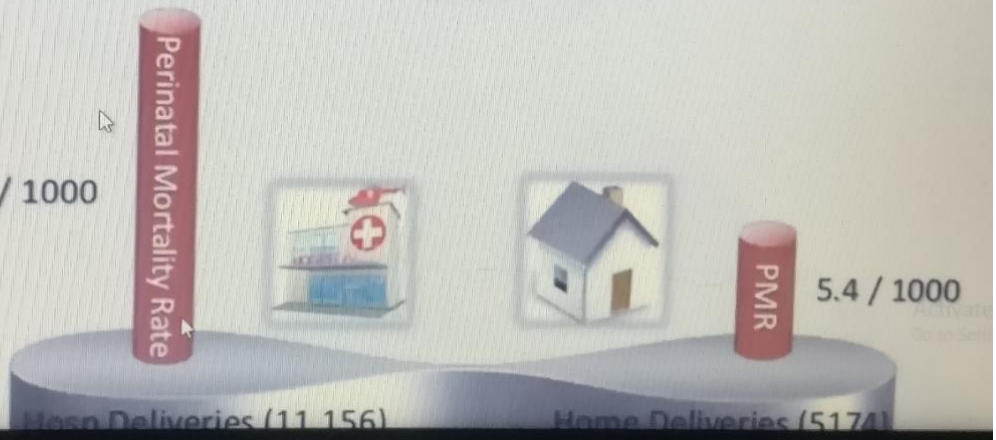
Positive correlation	Negative correlation
There is increase in value of dependent variable with increase of independent variable	There is decrease in value of dependent variable with increase of independent variable
Examples: •Water consumption and body temperature •Study time and grades of students	Examples: •Alcohol consumption and driving ability

TYPES OF ASSOCIATION

1. SPURIOUS ASSOCIATION
2. INDIRECT ASSOCIATION
3. DIRECT ASSOCIATION
   - A. One-to-One Causal Relationship
   - B. Multi-Factorial Causation

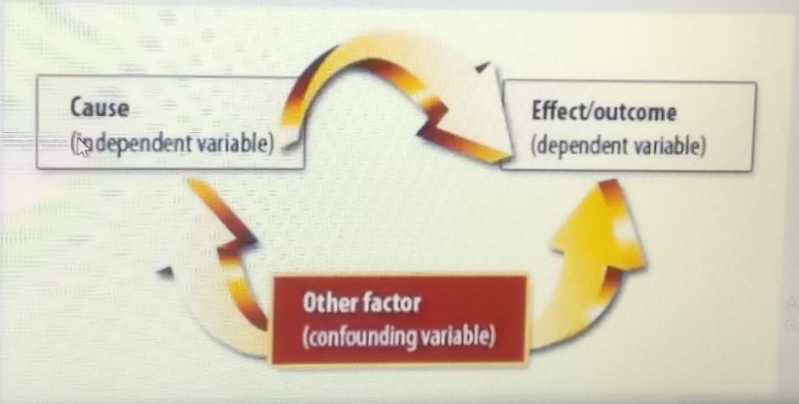
SPURIOUS ASSOCIATION

- ✓ Some observed associations b/n a suspected factor and disease may not be real.
- ✓ This Fallacy of presumption arises when two variables are improperly compared (due to Bias).



INDIRECT ASSOCIATION

- It is a statistical association between a characteristic of interest and a disease due to the presence of another factor i.e. common factor (*Confounding variable*).



Confounding Factor

1.

Confounder (e.g. smoking)

1. Smoking is a known cause of CVD

Disease outcome (e.g. CVD)

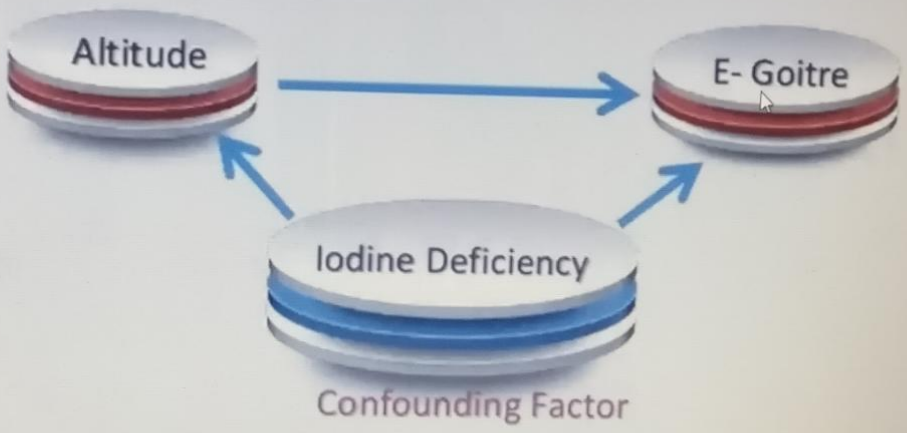
2. Coffee drinking could be common among smokers (not a causal link)

Factor (coffee drinking)

3. If so, drinking coffee will appear to be linked to CVD

McMohan Study (Pancreatic

2.



3. *Yudkin & Roddy's* wrong hypothesis on Sucrose and CHD association (Smoking is the Confounder)

4. *Jacob Yerushalamy* identified the association b/n Smoking and Low birth weight babies is due to Confounding.

DIRECT ASSOCIATION

A. One-to-One Causal Relationship

- This model suggests that two factors (A & B) exhibit one to one relationship, if – *Change in A is followed by Change in B.*

Paramyxo Virus

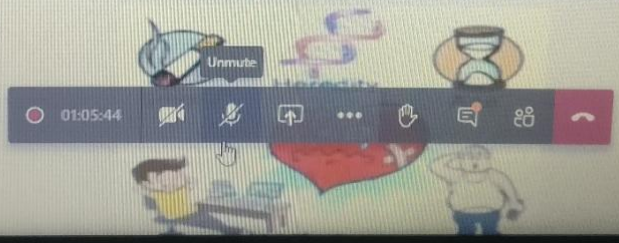
Measles

B. Multifactorial Causation

- In Several Modern Diseases, more than one factor is implicated in the Web of Causation.

  Eg: Both *Asbestos exposure* and *Smoking* cause Lung Cancer independently.

- As our Knowledge on disease increases, we may discover a common biochemical event, which can be altered by each of these factors



OBSERVED ASSOCIATION

- Could it be due to BIAS? — No
- Could it be CONFOUNDING? — No
- Could it be result of CHANCE? — No
- Could it be CAUSAL RELATION? — Yes
- Apply Guidelines and Make Judgement.