Epidemiological studies

Dr. Akhilesh Bhargava
MD, DHA, PGDHRM
Prof. Community Medicine
&
Director-SIHFW, Jaipur

Epidemiological studies

- **Descriptive**
  - Correlation studies
  - Individual studies
- **Analytical**
  - Case control studies
  - Cohort studies
- **Experimental**
  - Randomized design
    - Blind
    - Double blind
    - Triple blind
  - Clinical trials

Types of Descriptive studies

a. Population (Correlation) studies.

b. Individual studies

i. Case reports

ii. Case series

iii. Cross sectional studies (Prevalence studies)

Population (Correlation) studies.

Use data from entire population to compare-

- Disease frequency between *different population* groups during *same period*
- Disease frequency in *same population at different periods*.

useful in formulation of a hypothesis but not for testing a hypothesis.

Advantages and Limitations

- Inexpensive in terms of Time & Money
- Routinely available information can be used

- Exposure cannot be linked with disease as whole population is represented.
- Lack of ability to control the effects of potential confounding variables.
- Presence of a correlation does not necessarily mean a statistical association.
- Correlation data represent average exposure levels rather than actual individual levels

Individual studies

- Case reports (single patient)
- Case series (group of patients with similar diagnosis)

Individual studies: Advantages and Limitations

- Cases can be aggregated from different sources to generate hypotheses and describe syndromes
- Statistical association can not be established as there is no comparison group

Cross sectional studies

- Measure disease and exposure simultaneously in a defined population over a defined period
- provide instant information about frequency and characteristics of a disease
- Useful in-
  - Assessing health status
  - Identifying health care needs.
  - Providing data on Prevalence of Disease, disability and utilization of services
  - Provide data for Health care planning and administration

Advantages and Limitations

- Short term so less expensive
- Offer starting point in prospective studies
- Allow assessment of risk though less precise
- Start from a reference population from where sample is drawn, generalization can be made

- Not possible to ascertain whether exposure preceded or resulted from the disease.
- Since prevalence is assessed, results are affected by survival factors

Case-Control study

- A non-experimental Study,
- Subjects enrolled based on presence/absence of outcome,
- Cases/controls compared with regard to prior exposure to causal factors.

Designing: case-control

- Criteria-

  Comparability of cases & controls

- Issues –

  # Defining & selection of cases

  # Selection of controls

  # Information on disease exposure status

Case-control: strengths & limitations

- Study of diseases with long latent period
- Low cost
- Short time
- Applicable to rare diseases also

- Bias in selection/Reporting/Recording.

Case-control : selection of cases

- Requirements-
  - A standard case definition
  - Inclusion of likes, exclusion of dislikes
  - A strict diagnostic criteria for homogeneity
  - Only newly diagnosed cases be included- Older cases may distort the presentation
- Selection –
  - Hospital based
  - Pop. based

Case Control Studies

Group of interest eg.
Cancer patients

Take histories

Compare histories

Draw conclusions

Take histories

Comparison group eg.
Non-patients

Cases

- Selection criteria
- Sources
  - ✓ Gen. Population
  - ✓ Hospital Patients
    - o Whole group
    - o Sample

Controls

- Selection criteria
- Sources
  - ✓ Hospital (bias ?)
  - ✓ Relative (genetic)
  - ✓ Neighbors
  - ✓ Gen. population

Control selection criteria

- Characteristics and sources of Cases
- Need to obtain comparable information from cases and controls
- Practicality
- Cost
- Representative of population from where cases are drawn

Advantages/Disadvantage of control sources

Source	Advantage	Disadvantage
Hospital	- easy identification - belong to same class - cooperative - min. non-response	- do not represent exposure in general population - admission bias
Gen .Pop.	- high comparability	- costly, more time, - recall bias - loss during study

Selection of controls

- How many control groups ?
  - Hospital – more than one
  - Where one group has some deficiency (stress & peptic ulcer in executives) and stress present in hospitalized
- How many subjects in a control group?
  - 1:1 if no. of cases & controls is large and cost of getting information is same
  - Max. 4:1, more does not increase statistical strength but cost increases

Case Control Study Approach

Status	Outcome	No Outcome
Exposed	a	b
Non-exposed	c	d

$$\text{Odds Ratio} = \frac{\text{Odds of exposure in cases}}{\text{Odds of exposure in controls}} = \frac{a / c}{b / d} = \frac{a\,d}{b\,c}$$

Odds Ratio = Relative Risk Conditions & Interpretations

OR –an unbiased and valid estimate of RR in Case control studies, as-

only newly diagnosed are included
selection is not based on exposure level

An OR=1- no statistical association

- exposure is not a risk factor

OR >1 – a positive association between exposure and outcome

OR<1 – Less risk, or even protective value of a Risk factor.

Bias and its play in Case Control study-

Bias

"any systematic error in the study that results in an incorrect estimate of the association between exposure and risk of disease".

Types of Bias-

1. Selection Bias
   a. Prevalence-Incidence Bias
   b. Admission rate (Berkson's) Bias
   c. Non-response/ Refusal Bias

2. Observational or Information Bias
   a. Diagnostic Bias
   b. Recall Bias

How Bias gets in

- Differential surveillance,
- Diagnosis & Referral during ascertaining status of subjects,

  OR ,

  for-Unwillingness to participate ,

- Non selection by investigator &
- Replacement of originally selected
- Selection bias- Refusal
  Non-response
  Self selection

Cohort

group with common experience

Roman Army Groups

Cohort study-

"A non-experimental study, subjects enrolled based on exposure level to main independent variable, followed to determine development of the dependent variable"

Cohort study

Characteristics-

- Absence of disease at initiation point
- Follow up over a period of time

Types-

Retrospective (after exposure but certainly before disease/ outcome)

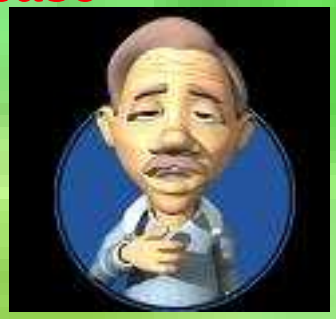
Prospective (exposure may/may not have but outcome certainly not)

Case control study

Exposure **Disease**

?----------

?----------

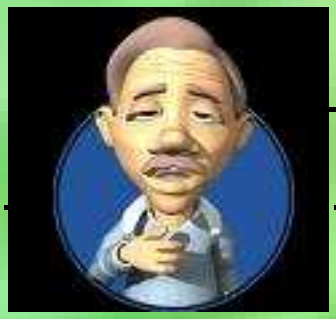


Prospective Cohort Study

Exposure **Disease**

----------?

---------- ?

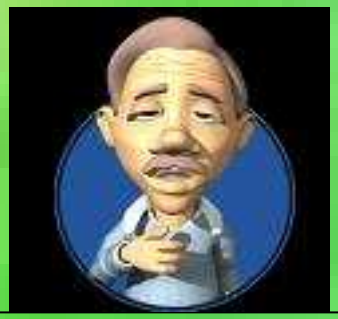


Retrospective cohort Study

Exposure **Disease**

---------- ?

---------- ?



Study design - considerations

General- Hypothesis to be tested
Resources
Current state of knowledge

Which of cohort study ?
Time
Money
Latent pd. of disease
Availability of information/Records
Frequency of occurrence of disease
Sample size
Follow up period required

Issues in Cohort study design

Selection –

- Exposed group
- Comparison group

Sources of data –

- Exposure data
- Outcome data

Approaches to follow-up

Cohort study...
Selection-Exposed/Comparison groups

Exposed

Criteria- Availability

Complete & accurate information

Exposure

Comparison

Criteria- Similarity but for exposure

Cohort...
Sources of information

Information- complete & accurate

Exposed- Gen. population

Special exposure groups

Comparison- Graded exposure groups

Individuals from same work cohort without exposure to risk factor

Rates of disease in gen. pop.

Gen. Population

Cohort...
Sources of exposure information

- Medical records
- Employment records
- Insurance records
- Study subject's interrogation
- Direct measurement of risk in environment

Cohort...
Sources of outcome data

Depend upon- specific resources
disease under evaluation

Objective-- complete, comparable &
unbiased data

Sources- Death certificates
Periodic examinations
Hospital discharge records/tickets
Study participants
Relatives
Insurance records

Cohort...Follow up

p e r i o d of t i m e

Exposure----------Outcome

BIAS ← **change in composition of group**

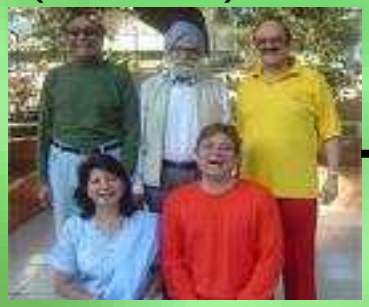
↓

movement

change in job/names/ residence/habits

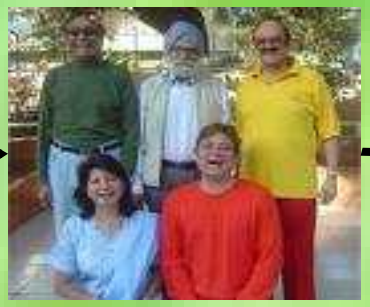
Cohort Studies

Group of interest
(smokers)



Follow

Over time



Compare
outcomes

Control group
(non-smokers)

Follow

Over time

Cohort study

Strengths

- Can measure RR/AR
- Handle rare diseases
- Establish temporality
- Easy identification of exposed group
- Minimal Bias
- Multiple effects of same exposure

Weaknesses

- Time/cost-expensive
- High follow-up loss –
  - changes in
    - composition
    - size
    - environment
    - habits
    - interest
- Large sample-manageable?

Cohort...
Issues in analysis/interpretation

- **Role of Bias-**

  (More in Retrospective than Prospective)

  Determinants: classification of subjects

  ascertainment of outcome

- **Losses due to follow-up**
- **Effect of non-participation**

  Participation dependent on :awareness levels

  motivation

  exposure to risk

Cohort Study Approach

Exposure	Outcome (Disease)	
	Yes	NO
Present	A	B
Absent	C	D

Relative Risk = Incidence in exposed / Incidence in un-exposed = A / A+B / C / C+D

Cohort Study: Outbreak of Gastroenteritis associated with eating Cheese

	AGE	NO AGE	
Ate	700 (A)	140 (B)	840 (A+B)
Did not eat	30 (C)	400 (D)	430 (C+D)

$$RR = \frac{700 / 840}{30 / 430} = 11.9\ \%$$

Exposure	Outcome		
	Present	Absent	
Smoker	120 (A)	280 (B)	400
Non-smoker	30 (C)	570 (D)	600

OR= AD/BC=120 X 570/280 X 30= 8.14

$$RR = \frac{A/A+B}{C/C+D} = \frac{120/400}{30/600} = 6.0$$

AR= RR-1/RR =6-1/6 =0.83

Attribute	Cohort	Case-control	Cross-sectional
Pop.	Disease free	Cases& control	Pop. with no dis-. no exp., no dis.- but exp., no exp- with dis., exp.-with dis.
Sample	healthy	unknown source pop. for cases	survivors at a pt.
Temporal sequence	Pros/retro.	Retrospective	Retrospective
Function	Compares incidence	Prevalence	Describes association
Outcome Measure	Incidence RR/AR	Prevalence OR	Prevalence OR
Causality	Strong	needs careful analysis	only suggestive
Bias	Manageable	can be	Difficult

Basis	Cohort	Case-control	Cross-sectional
Rare dis.	Not practical	Best	NA
Determine risk	Best	Only estimate	Prevalence
Whether exp. Preceded	Best	NA	NA
Time/money	Expensive	least expensive	less expensive
Planning	long term	NA	Best