Cohort study

Learning objectives

At the end of this lecture, students will be able to:

1. define cohort, cohort study, and certain basic epidemiological terms
2. organize a cohort study
3. describe selection of the study groups
4. demonstrate the advantages and disadvantages of cohort studies

WHAT IS COHORT?

''*cohors*' (Latin word) = Refers to warriors and gives notion of a group of persons proceeding or banded together in time. Epidemiologically refers to a group of persons with a common statistical characteristic. e.g. age, birth date, smoking ... etc.

Definition & Synonyms

The cohort study is an observational analytic epidemiological study, which attempts to study the relationship (association) between a proposed risk factor (exposure) and the subsequent risk of developing disease.

Synonyms

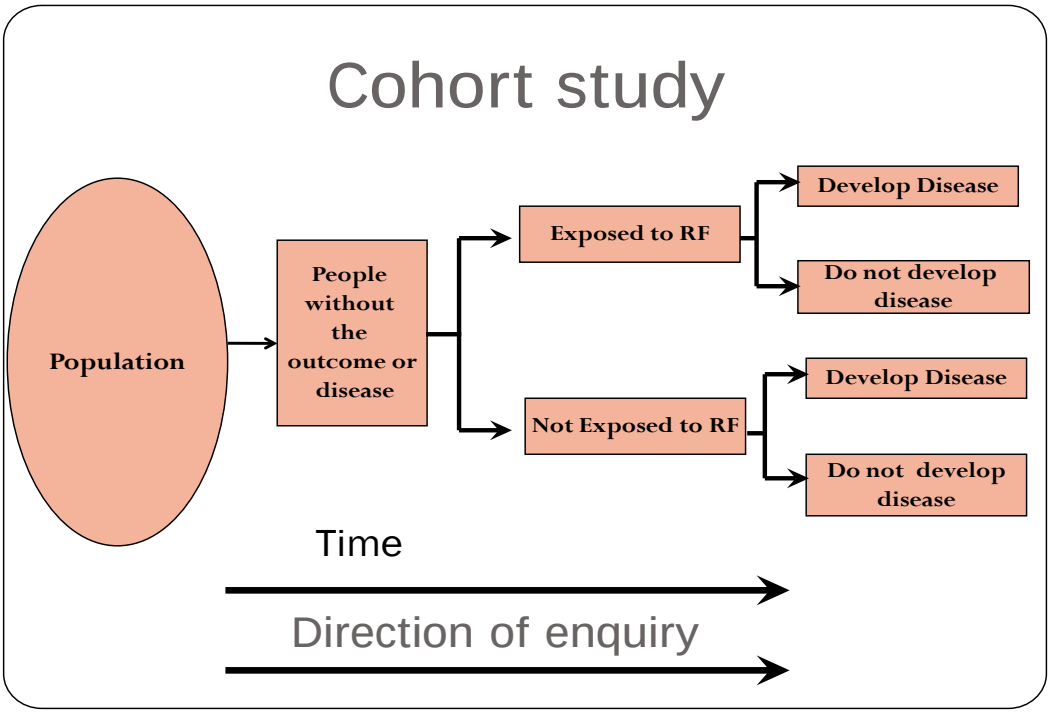
- Follow-up
- Longitudinal
- Prospective
- Incidence study

Indications for a cohort study

- When there is good evidence of an association between exposure and disease, established by any observational study
- When exposure is rare but incidence of disease is higher among exposed
- When follow-up is easy, and cohort is stable
- When sufficient funds are available

The cohort design

- starts with people free of disease
- assesses exposure at “baseline”
- assesses disease status at “follow-up”
- A definite beginning and end



Elements of cohort study (Steps in conducting cohort study)

1. Selection of the cohort groups
    a. Cohort study subjects (Group with exposure)
    b. Comparison group (Group without exposure)
2. Obtaining data on exposure
3. Follow up
4. Data analysis

1. Selection of the cohort groups

General consideration while selection of cohorts

Both the cohort study (exposed) group and controls (non-exposed) are free of the disease
Both groups should be equally susceptible to disease
Both groups should be comparable
Diagnostic criteria for the disease should be defined well earlier

A. Selection of cohort study subjects (Group with exposure):

The sources of the study subjects

1) General population
   - Whole population in an area
   - A representative sample
2) Special group of population
   - Selected group such as Occupation group / professional group
   - Exposure groups: such as persons having exposure to some physical, chemical or biological agents. e.g. X-ray exposure (radiologists)

B. Selection of comparison group (Group without exposure

1) Internal comparison

Only one cohort is involved in study (e.g. smokers), which is sub classified and internal comparison done (e.g. Intensity of smoking)

2) External comparison

- More than one group in the study for the purpose of comparison (study group and control group)
- e.g. Cohort of radiologists compared with ophthalmologists

3) Comparison with general population rates

- If no comparison group is available, we can compare the rates of study cohort with general population. e.g. we compare cancer rate in uranium miners with cancer rate in general population.

2. Obtaining data on exposure

- Personal interviews / mailed questionnaires
- Reviews of records: such as dose of drug, radiation, type of surgery ...etc
- Medical examination or special test: e.g. Blood pressure, serum cholesterol
- Environmental survey

By obtaining the data of exposure, we can classify cohorts as:

- - Exposed and non exposed, or
- - By degree of exposure, we can sub-classify cohorts into: Mild, moderate, and high exposure

3. Follow-up

- To obtain data about outcome to be determined (morbidity or death), we have to use any one way of the following:
  - - Mailed questionnaires, telephone calls, personal interviews
  - - Periodic medical examination
  - - Reviewing records
  - - Surveillance of death records
  - - Follow up is the most critical part of the study
- Some loss to follow up is inevitable due to death, change of address, migration, or change of occupation.
- Loss to follow-up is one of the drawbacks of the cohort study.

4. Analysis of data

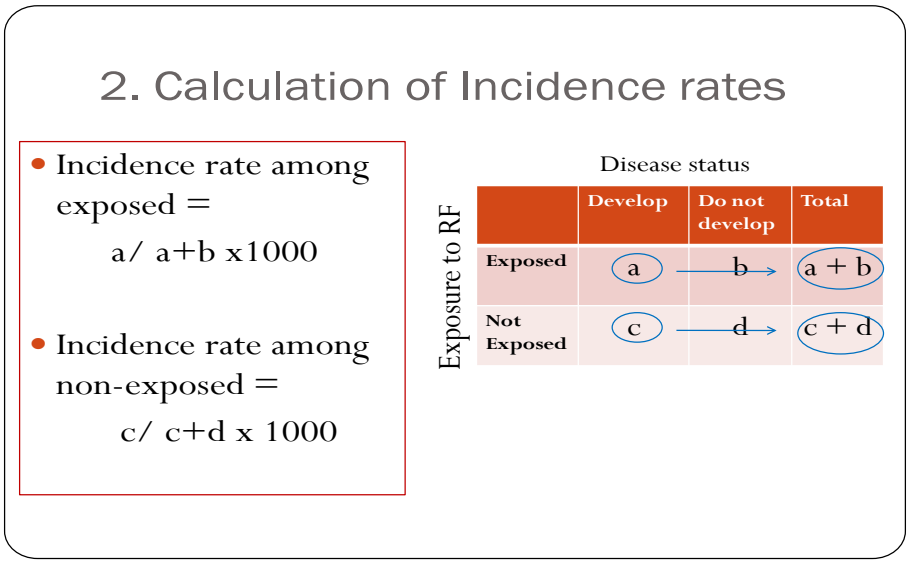
- **A.** Arrangement of data in 2x2 table

1. Standard 2 x 2 table

		Disease status: Develop	Disease status: Do not develop	Total
Exposure to risk factor	Exposed	a	b	a + b
	Non-Exposed	c	d	c + d

-
-
-

- **B.** Calculation of incidence rates among exposed and non exposed groups



C. Estimation of Relative risk

Relative risk is a measure of strength of association between an exposure (risk factor) and an outcome (disease). Its value is an indicator of the significance of the exposure in the aetiology of the outcome. It is calculated according to the following formula:

$$RR = \frac{\text{Incidence rate of disease among exposed } (a/a+b)}{\text{Incidence rate of disease among non-exposed } (c/c+d)}$$

$$RR = \frac{a/a+b}{c/c+d}$$

Interpretation of Relative Risk (RR)

When **RR=1**: it means no association between exposure and disease. The incidence rates are identical between groups

When **RR> 1:** this means a positive association (increased risk**)** i.e. the exposed group has higher incidence than the unexposed group

When **RR< 1**: indicates a negative association (protective effect) i.e. the unexposed group has higher incidence than exposed group

Guide to the strength of an epidemiological association

Guide to the strength of an epidemiological association

Relative risk	strength of association
1.0	None
> 1.0 - < 1.5	Weak
1.5–3.0	Moderate
3.1–10.0	Strong
> 10	Infinite

D. Estimation of Attributable Risk (AR): It is the fraction of risk that can be attributed to the exposure risk factor under study.

Attributable Risk = IR of disease among exposed – IR disease among non exposed

E. Attributable risk percent [Percentage of reduction]

It refers to the proportion of the total risk of disease that can be reduced if the exposure to risk factor is eliminated.

Also, we can calculate AR% (Exposed) **[Percentage of reduction in risk if the risk factor is removed]**

$$AR\%\ (\text{percentage of reduction}) = \frac{[IR\ (Exposed) - IR\ (Unexposed)]}{IR\ (Exp)} \times 100$$

Example of analysis of data:

Find out RR and AR for the following data

Smoking	Hypertension		Total
	YES	NO	
YES	8	192	200
NO	4	196	200
	12	388	400

IR(exposed)=8/200 x 1000= 40/1000

IR(non-exposed=4/200 x 1000=20/1000

RR= 40/20= 2

AR= 40-20=20/1000

AR% (Percent of reduction)= 40-20/40 x100= 50%

Advantages of cohort study

- Can measure incidence and risks
- An efficient method for studying rare exposure
- Assesses multiple outcomes of a single exposure
- Establishes temporal relationship between exposure and outcome (Exposure happened before outcome)
- Low potential for bias [Avoids recall bias]
- Does not require strict random assignments of subjects
- ***Cohort study is the best observational design to establish association between risk factor and an outcome***

Disadvantages

- Expensive and time-consuming
- Loss to follow-up
- Bias in ascertainment of exposure
- Measurement errors, multiple interviews, tests
- Complexity of data analysis

Questions:

1. In a population of 10500 adult persons, the man: woman ratio was 1.1:1. Both men and women were followed–up for 2 years.1650 men and 20% of women developed hypertension during the follow-up period.

A. Is there an association between gender and hypertension? Do appropriate calculations to prove your answer.
B. What is the overall annual incidence rate of hypertension in that population?
C. If men and women have the same risk of development of hypertension, what will be the expected percentage of reduction in risk of hypertension in the exposed group?

2. In a cohort study the relative risk for chronic obstructive pulmonary disease (COPD) for moderate smokers versus non-smokers was 4. For heavy smokers compared to non-smokers the relative risk was 10. What would have been the relative risk for COPD in this study if the heavy smokers were used as the reference category?

A. For non-smoking 0.1 and for moderate smoking 0.4
B. For non-smoking 0.2 and for moderate smoking 0.6
C. For non-smoking 4 and for heavy smoking 10
D. For non-smoking 0.4 and for moderate smoking 1
E. This cannot be calculated with the available data