Cohort Study

Concept of cohort

The word cohort is derived from the Latin “cohorts” meaning an enclosure, company, or crowd.

In Roman times a cohort was a body of 300–600 infantry.

In epidemiological terms the cohort is a group of people with something in common, usually an exposure or involvement in a defined population group.

Definition

- Cohort study is a type of analytical study which is undertaken to obtain additional evidence to refute or support existence of association between suspected cause and diseases.

- Other names of cohort study are Longitudinal study, Incidence study and forward looking study

Features of cohort studies

- Cohorts are identified prior to appearance of disease under investigation
- The study groups are observed over a period of time to determine the frequency of disease among them
- The study proceeds from cause to effects

Indications for cohort study

- There is good evidence of an association between exposure and disease, from other studies.
- Exposure is rare.
- Attrition of study population can be minimized.
- Sufficient fund is available.

Framework of cohort study

Design of Cohort Study

First, identify

			Totals	
Exposed			a + b	
Not exposed			c + d	

Then

Then, follow to see whether

Calculate and compare

	Disease develops	Disease does not develop	Totals	Incidence of disease
Exposed	a	b	a + b	$\frac{a}{a+b}$
Not exposed	c	d	c + d	$\frac{c}{c+d}$

(a+b) is called study cohort and (c+d) is called control cohort

Consideration during selection of Cohort

- The cohort must be free from disease under study.
- Insofar as the knowledge permits, both the groups should be equally susceptible to disease under study.
- Both the groups must be comparable in respect of all variable which influence the occurrence of disease
- Diagnostic and eligibility criteria of the disease must be defined beforehand.

Types of cohort study

- Prospective study
- Retrospective cohort study

Prospective cohort study

- The common strategy of cohort studies is to start with a reference population (or a representative sample thereof), some of whom have certain characteristics or attributes relevant to the study (exposed group), with others who do not have those characteristics (unexposed group).

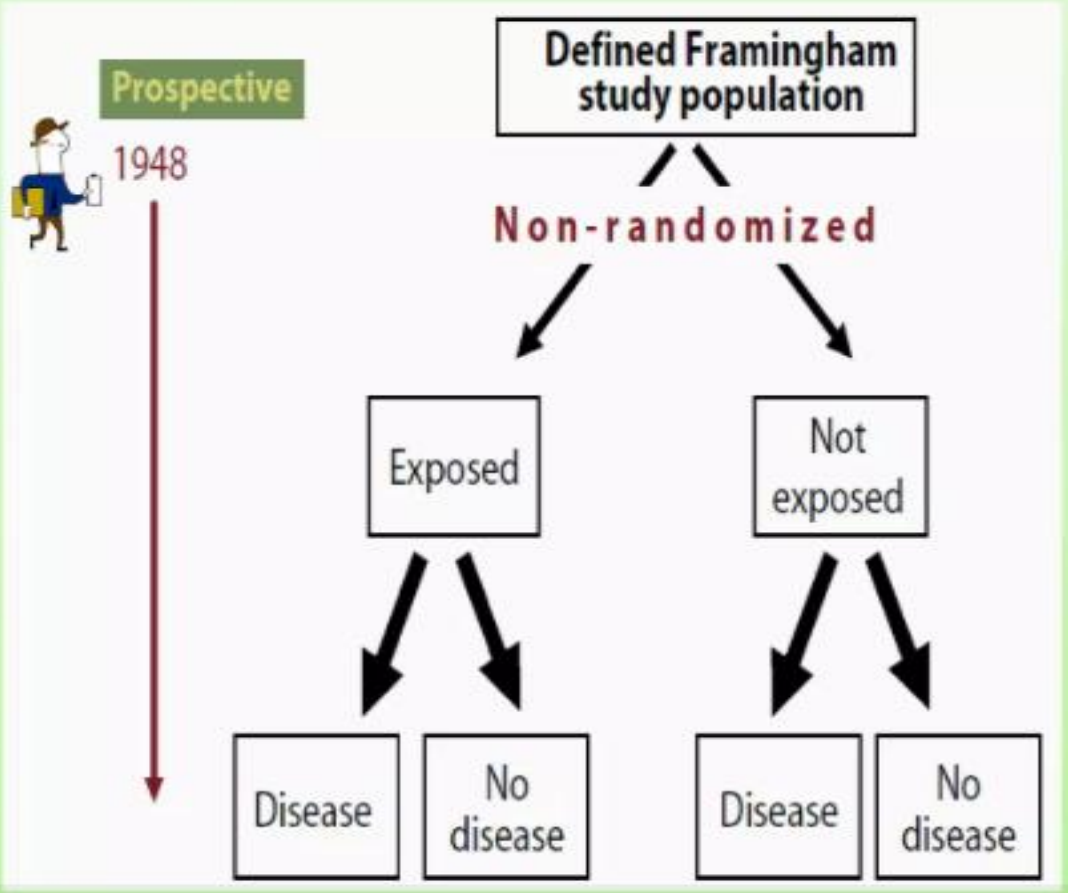
- Both groups should, at the outset of the study, be free from the condition under consideration. Both groups are then observed over a specified period to find out the risk each group has of developing the condition(s) of interest.

Example of Prospective Cohort Study

Framingham Heart Study

- Initiated in 1948 to study the relationship of a variety of factors to the subsequent development of heart disease with 5127 samples(30 to 59 yrs) at Framingham.
- Study subjects were examined every 2 yrs for 20 years.
- Daily Surveillance of hospitalization at Framingham hospital.
- Study found that Hypertensive, tobacco smoking, elevated blood cholesterol are associated to CHD
- Increased physical activity associated with decreased risk of CHD

Framework



Problem of prospective study

- Study might take long duration.
- Sufficient amount of funding for long period.
- Missing of study subjects.

Retrospective Cohort Study

- A retrospective cohort study is one in which the outcome have all occurred before the start of investigation.
- Investigator goes back to the past to select study group from existing records of the past employment, medical and other records and traces them forward through time from the past date fixed on the records usually to the present.
- Known with the name of Historical Cohort and noncurrent cohort

Example of Retrospective Study

- Suppose that we began our study on association between smoking habit and lung cancer in 2008
- Now we find that an old roster of elementary schoolchildren from 1988 is available in our community, and that they had been surveyed regarding their smoking habits in 1998.
- Using these data resources in 2008, we can begin to determine who in this population has developed lung cancer and who has not.

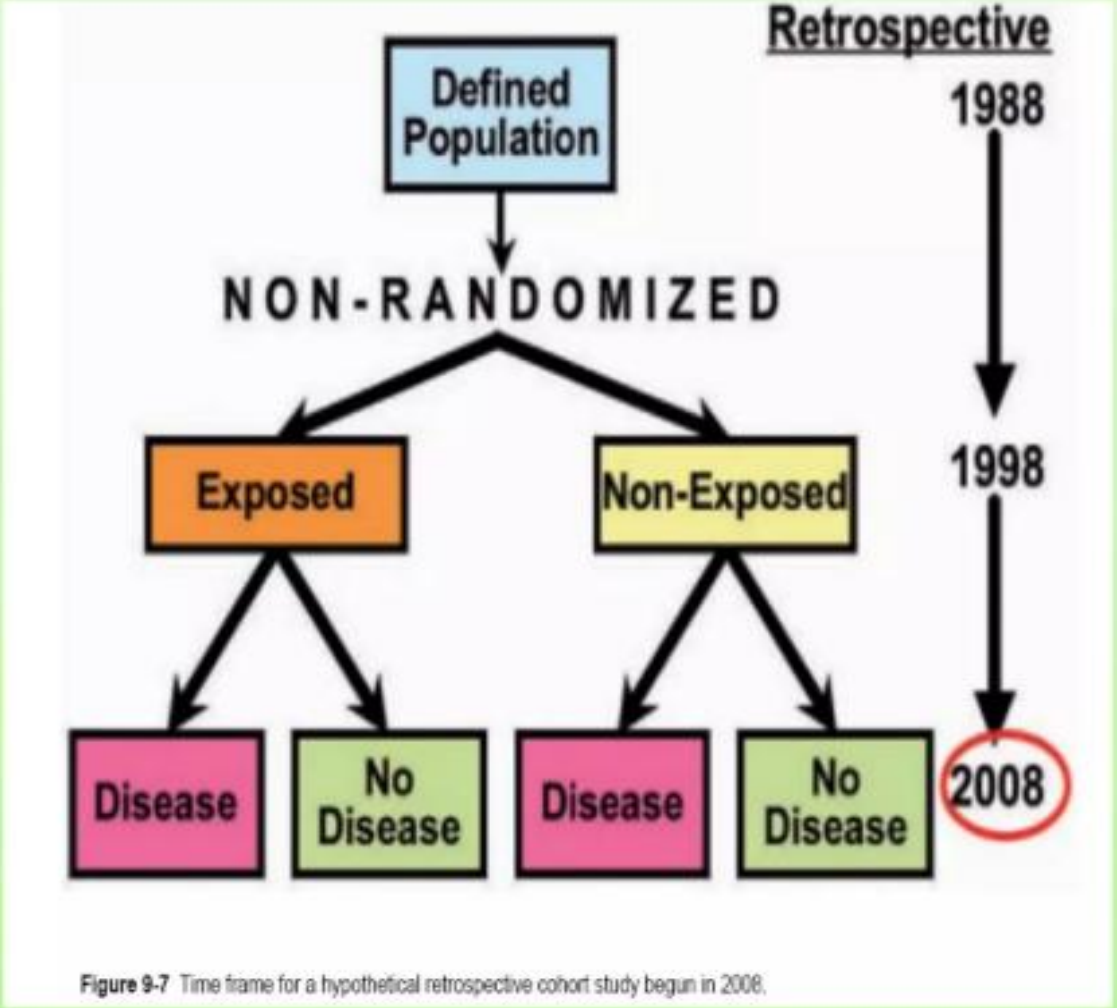


Figure 9-7 Time frame for a hypothetical retrospective cohort study begun in 2008.

Comparison of retrospective and prospective cohort study

Attribute	Retrospective approach	Prospective approach
Information	< complete < accurate	> complete > acurate
Emerging new exposures	Not useful	Useful
Expense	Less costly	More costly
Completion time	Shorter	Longer

Steps of Cohort Study

1. Selection of study population
2. Obtaining Data on exposure
3. Selection of comparison group
4. Follow up
5. Analysis

1. Selection of study subjects

The usual procedure is to locate or identify the cohort, which may be a total population in an area or sample thereof. Cohort can be:

- community cohort of specific age and sex;
- exposure cohort e.g. radiologists, smokers, users of oral contraceptives;
- birth cohort e.g. school entrants;
- occupational cohort e.g. miners, military personnel;
- marriage cohort;
- diagnosed or treated cohort, e.g. cases treated with radiotherapy, surgery, hormonal treatment.

2. Obtaining data on Exposure

- From Cohort Members : Personal interview, mailed questionnaire
- Review of Records : Certain kinds of information like dose of radiation, kinds of surgery received can only be obtained from medical records.
- Medical examination/ Special tests: In some cases information needs to be obtained from medical examination like in case of blood pressure, serum cholesterol,
- Environmental Survey of location where cohort lives

Information should be collected in a manner that allows classification of cohort according to

- whether or not they have been exposed to suspected factor
- According to level or degree of exposure
- Demographic variables which might influence frequency of disease under investigation

3. Comparison Group

Internal Comparison Group :

Single Cohort enters the study and its members on the basis of information obtained , can be classified into several comparison according to degree of exposure

Classification of exposure	No. of Deaths	Death rate
½ pack	24	95.2
½ to 1 pack	84	107.82
1-2 pack	90	229.2
+ 2 pack	97	264.2

Age Standardized death rate among 100000 men per year according to amount of cigarette smoking

External Comparison Group: when information on degree of exposure is not available.

if all workers at the factory had some degree of exposure, we would need to select a comparison group from another population, possibly another type of factory

Comparison with general population can also be used as comparison group

4. Follow UP

- The length of follow-up that is needed for some studies to reach a satisfactory end-point, when a large enough proportion of the participants have reached an outcome, may be many years or even decades.

- At the start of study, method should be determined depending on the outcome of study to obtain data for assessing outcome.

Procedure may be:

- Periodic medical examination of each member of cohort
- Reviewing physician and hospital records
- Routine surveillance of death records
- Mailed questionnaire, telephone calls and periodic home visits

5. Analysis

Data analyzed in terms of

- Incidence rate of outcome among exposed and non exposed

- Estimation of risk

Incidence rate

Choice between cumulative incidence and Incidence Density is a crucial issue

- Cumulative incidence: In cohort studies on acute diseases with short induction periods and a short time of follow-up, like outbreaks, the risk of disease can be estimated directly using the cumulative incidence, given a fixed cohort with fixed period of follow-up and a low fraction of drop-outs.

- Incidence Density: In cohort studies on chronic diseases with their long follow-up periods, however, the use of the cumulative incidence is not appropriate because usually disease-free follow-up periods differ strongly among cohort members. In such case incidence density is apposite measure

ANALYSIS OF COHORT STUDIES

	Outcome*			
	Death	No death	Incidence rate	Total
Exposed	A	B	A/(A+B)	A + B
Unexposed	C	D	C/(C+D)	C + D
Total	A + C	B + D		A+B+C+D

* Outcome : death/disease

A = Exposed persons who later develop disease or die
B = Exposed persons who do not develop diseases or die
C = Unexposed persons who later develop disease or die
D = Unexposed persons who do not develop diseases or die

The total number of exposed persons = A + B
The total number of unexposed persons = C + D

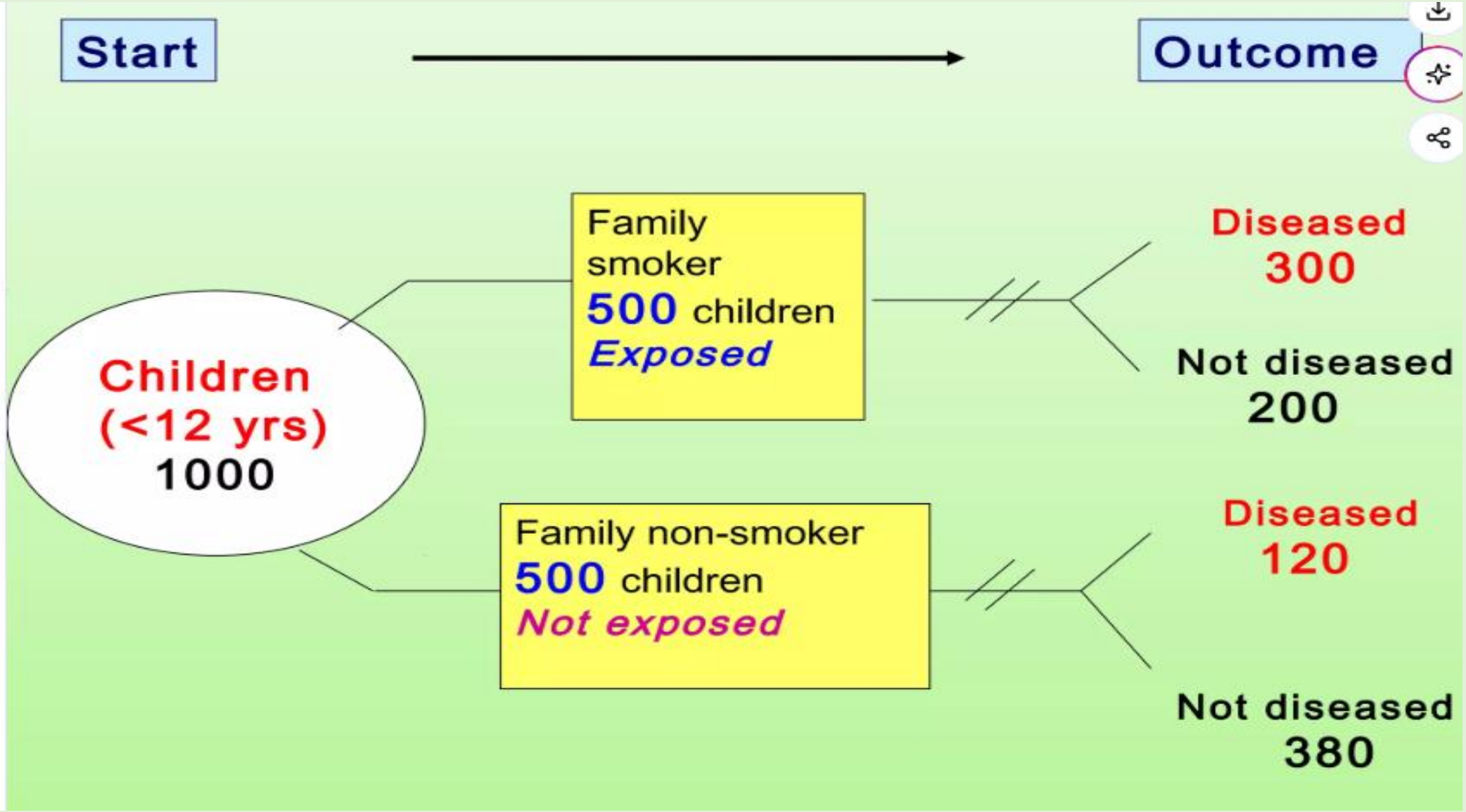
Incidence of disease(or death) among exposed= A/A+B
Incidence of disease(or death) among non-exposed= C/C+D

Relative Risk (RR)

- Estimates the magnitude of an association between exposure and disease
- Indicates the likelihood of developing the disease in the exposed group relative to those who are not exposed
- Ratio of risk of disease in exposed to the risk of disease in nonexposed

Relative Risk

$$RR = \frac{\text{Risk in exposed(Incidence in exposed group)}}{\text{Risk in non exposed(Incidence in non exposed group)}}$$


Start
Outcome
Children
(<12 yrs)
1000
Family
smoker
500 children
Exposed
Diseased
300
Not diseased
200
Family non-smoker
500 children
Not exposed
Diseased
120
Not diseased
380

Rate: Incidence rate

- Incidence of Resp. Infection among exposed children:

$$\frac{300}{500} = 60\%$$

- Incidence of Resp. Infect. Among non exposed children:

$$\frac{120}{500} = 24\%$$

Relative Risk: $\frac{\text{Incidence rate among exposed}}{\text{Incidence rate in non exposed.}}$

Risk Ratio

$$\frac{60}{24} = 2.5$$

Exposed individuals are 2.5 times more likely to develop disease than non exposed individuals.

Difference Measures

- **Attributable risk**
  - No. of cases among the exposed that could be eliminated if the exposure were removed

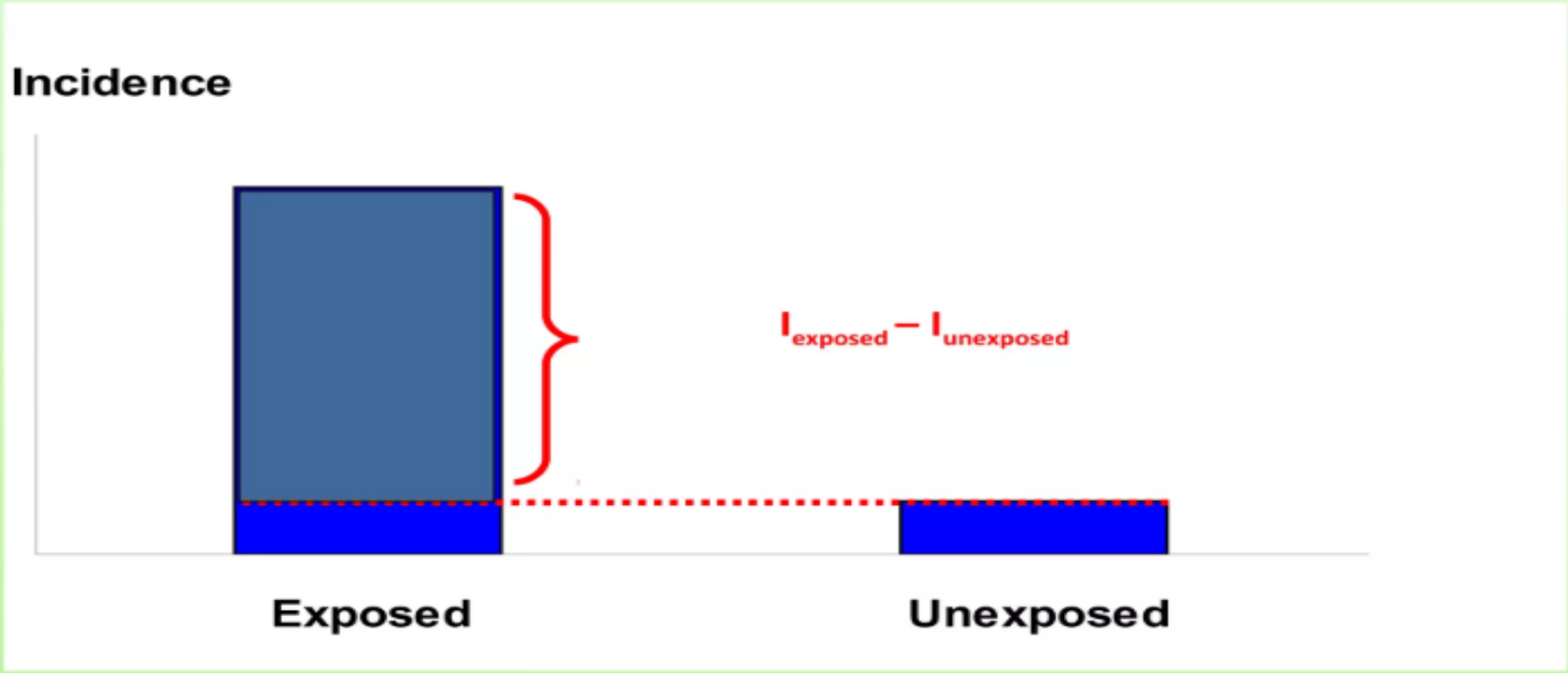
  = Incidence in exposed - Incidence in unexposed

- **Population Attributable Risk percentage:**

PAR expressed as a percentage of total risk in population

$$PAR\% = \frac{I_{population} - I_{unexposed}}{I_{population}} \times 100$$

Attributable Risk



I = Incidence

AR: Smoking and Lung cancer

	Lung Cancer				
Smoking	Yes	No		Incidence	RD
Yes	100	1900	2000	0.05	0.04
No	80	7920	8000	0.01	
	180	9820	10000		

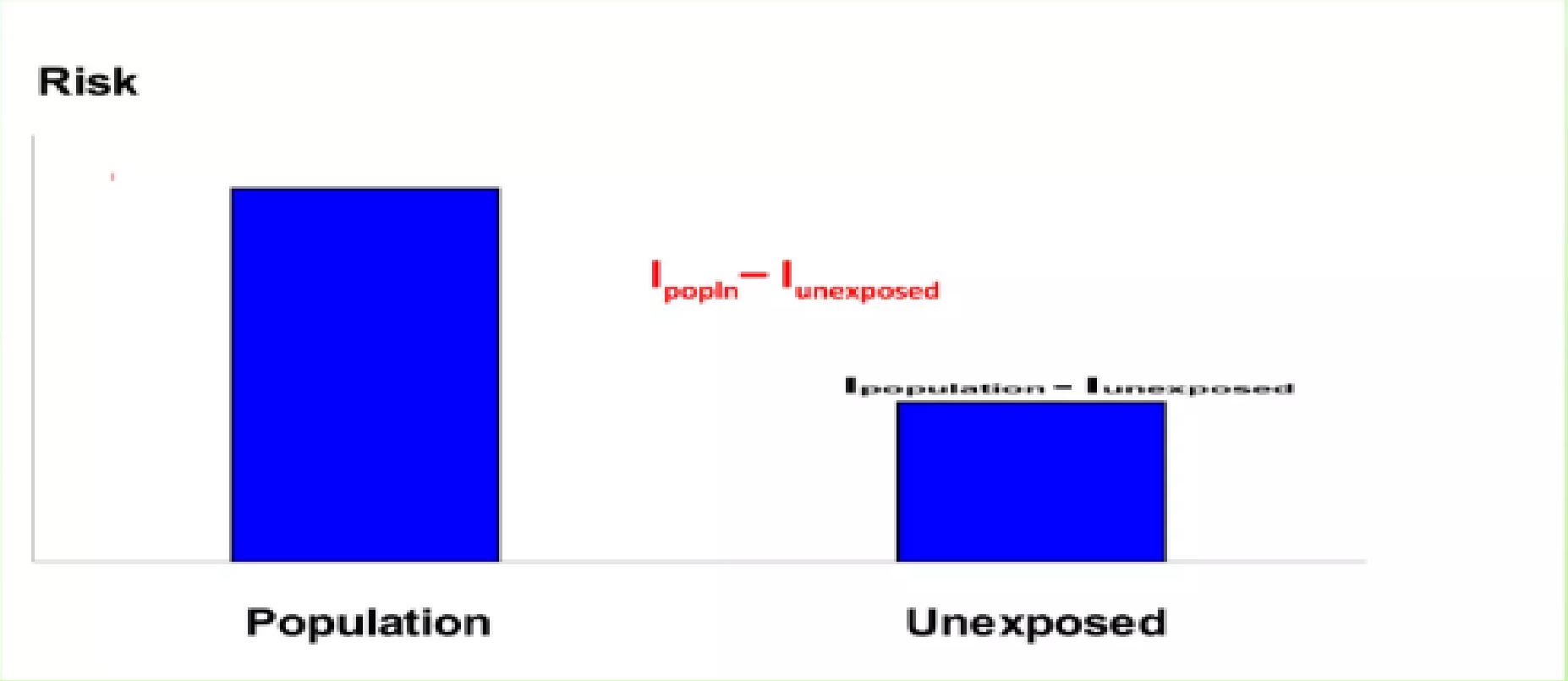
Attributable risk = **Incidence in exposed - Incidence in unexposed**
=0.5-0.1
=0.4

Population Attributable Risk (PAR)

- Excess risk of disease in total population attributable to exposure
- Reduction in risk which would be achieved if population entirely unexposed
- Helps determining which exposures relevant to public health in community

$$PAR = I_{population} - I_{unexposed}$$

Population Attributable Risk



PAR: Smoking

	Lung Cancer			
Smoking	**Yes**	**No**		**Risk**
Yes	**100**	**1900**	**2000**	**Incidence in exposed= 0.050**
No	**80**	**7920**	**8000**	**Incidence in unexposed=0.010**
	180	**9820**	**10000**	**Incidence in population=0.018**

$$PAR = 0.018 - 0.010 = 0.008$$

$$PAR\% = \frac{0.018 - 0.010}{0.018} \times 100 = 44\%$$

Conclusion:

44% of lung cancer in the population could be prevented if use of smoking were eliminated

Advantage of Cohort Studies

- Temporality can be established
- Incidence ca be calculated.
- Several possible outcome related to exposure can be studied simultaneously.
- Provide direct estimate of risk.
- Since comparison groups are formed before disease develops certain forms of bias can be minimized like misclassification bias.
- Allows the conclusion of cause effect relationship

Disadvantage of Cohort Studies

- Large population is needed
- Not suitable for rare diseases.
- It is time consuming and expensive
- Certain administrative problems like loss of staff, loss of funding and extensive record keeping are common.
- Problem of attrition of initial cohort is common
- Study itself may alter people's behavior