SCIENTIFIC MEDICAL RESEARCH

WRITER: إيهاب الرعود، أحمد سامي، أنس عنانزه

LECTURE: 6 + 7 + 8 + 9

Chapter 7: The Health Research Process

7.1 Types of Study Approaches (Designs):

There are many valid study approaches (8 are highlighted in your book):

** The design selected must be appropriate to the study goals.

FIGURE 7-1 Summary of Study Approaches

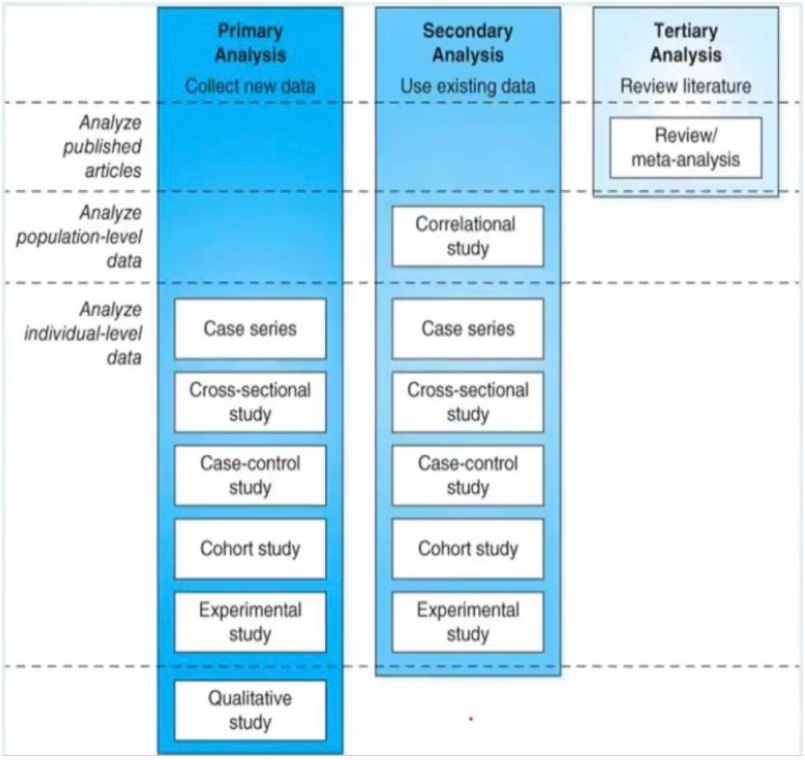
Study Approach	Goal
Case series	Describe a group of individuals with a disease
Cross-sectional survey	Describe exposure and/or disease status in a population
Case-control study	Compare exposure histories in people with disease (cases) and people without diseases (controls)
Cohort study	Compare rates of new (incident) disease in people with different exposure histories or follow a population forward in time to look for incident diseases
Experimental study	Compare outcomes in participants assigned to an intervention or control group
Qualitative study	Seek to understand how individuals and communities perceive and make sense of the world and their experiences
Correlational (ecological) study	Compare average levels of exposure and disease in several populations
Review/meta-analysis	Synthesize existing knowledge

For Example

If the goal is to
1. See weather an intervention is effective → experimental design
2. Understand population, describe patterns or to ask research question that are not focused on causitly → cross- sectional or cohort study

7.2 Primary, Secondary, Tertiary Studies:

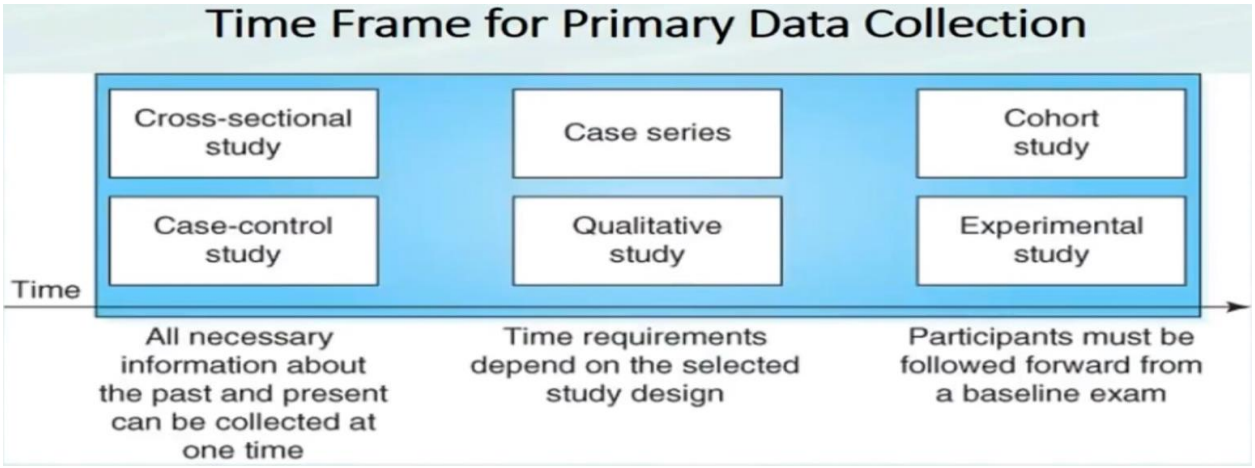
The first critical decision is whether to conduct a primary, secondary, or tertiary study. These are the study approaches, whereas what are included in the previous table are study designs (design is more specific than approach, but they can be used interchangeably). Now let's put them together >>>



Advantage of
* primary analysis → the researcher control over important details
* secondary and tertiary → the researcher may be able to move quickly from the definition of the study question to the analysis of related data

7.3 Study Duration: The time required for collecting and analyzing data varies from study to study.

Primary studies usually take longer than secondary or tertiary studies.



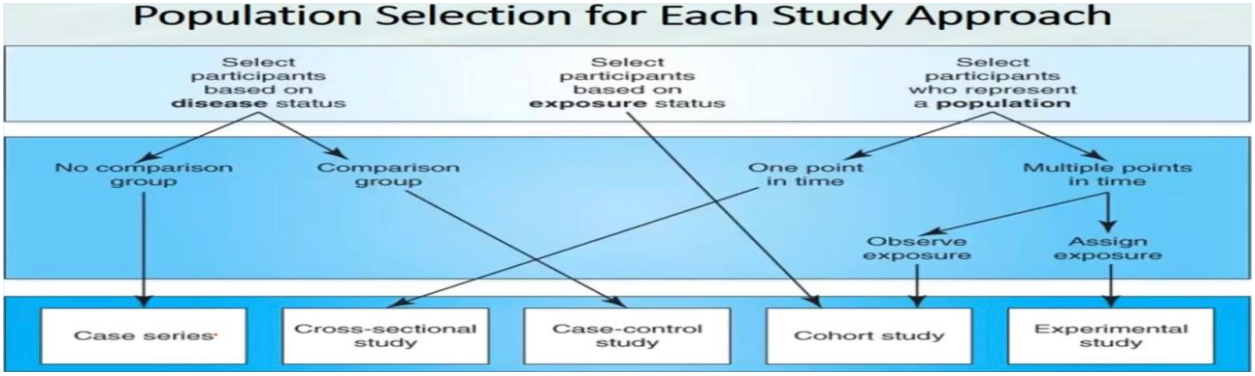
The more you go in time axis, the more effort, money, and time you need to these designs. And we will discuss each of them in details later…

Example
Secondary study → might be very short if an entire data file and the relevant supporting documentation can be downloaded from a website
old hospital chart

The duration of tertiary study → is high dependent on library access and on the number of publications that need to be acquired, read and summarized

7.4 Primary Focus: Exposure, Disease, or Population?

Primary study designs can be selected based on which EDPs is the major motivation for the study. The major decisions about which design to be selected are firstly determined by the study approach (primary/secondary/tertiary), then by your EDPs.



Case series>> When you want to describe a group of patients with certain disease, and this heavily relies on that they share a common diagnosis, for example.

Cohort study>> when you want to follow up a group of patients and check if certain exposure affects the rate or incidence of certain event or disease over time.

** based on population, you focus on a specific group of population according to age, sex, or certain health phenomena. We will discuss everything...

* Case series and case control → focused on individuals with a particular disease
* Cross sectional and some cohort study → seek to recruit population that is Representative of a well defined larger population.

8.1 Overview:

A case report *describes* one patient. Whereas a case series *describes* a group of individuals with the same disease or who have undergone the same procedure.

FIGURE 8-1 Key Characteristics of a Case Series

Objective	Describe a group of individuals with a disease
Primary study question	What are the key characteristics of the cases in this study population?
Population	All individuals in the study must have the same disease or be undergoing the same procedure.
When to use this approach	A source of cases is available, and no comparison group is required or available.
Requirement	An appropriate source of cases is available.
First steps	1. Specify what new and important information the analysis will provide. 2. Identify a source of cases. 3. Assign a case definition. 4. Select the characteristics of the study population that will be described.
What to watch out for	A lack of generalizability
Key statistical measure	Only descriptive statistics are required.

Both describe (description only) a uniqueness of disease itself or its occurrence (in another age or sex, for example).

Case definition: description of the case's characteristics.

Disadvantage: you can't generalize your findings.

8.2 Case Definitions

A clear case definition spells out inclusion & exclusion criteria.

ICD (International Classification of Diseases) **codes can be helpful.**

Code alone is rarely sufficient to cover all inclusion and exclusion criteria

Include person, place, and time (PPT) characteristics.

FIGURE 8-2 Sample Case Definitions

Category	Example 1	Example 2
Disease/ procedure	Whooping cough (ICD-10 code A37)	Liver transplantation
Person	Any person with a confirmed case of whooping cough, defined as an acute cough of any duration with isolation of Bordatella pertussis from a clinical specimen or a cough lasting 2 or more weeks with paroxysms of coughing, inspiratory "whoop," or posttussive vomiting and contact with a laboratory-confirmed case of pertussis	Adult patients (ages 18 and older at the time of transplant), excluding those who were not receiving their first liver transplant and those who received multi-organ transplants
Place	Residents of Big City whose diagnoses were reported to the Big City Health Department (which requires notification of all diagnoses of pertussis)	Patients who had transplant surgery at the Oakville Regional University Medical Center
Time	First sought clinical care between January 1 and March 31, 2016	Recipients of liver transplants between January 1, 2006, and December 31, 2014, who were followed for a minimum of 2 years post-transplant

In addition to that → case definition essential for any out break investigation, no matter study which approach is used to investigate the epidemic .

8.3 Special Considerations

Use a "questionnaire" (data collection sheet) to extract information from medical charts.

Remember that missing information doesn't mean that a symptom or sign was not present, just that it wasn't recorded in the file (variations among population).

Ethical approval is acquired, and great care must be taken to protect the identities of study subjects.

Photographs can only be used with written permission from the patient.

8.4 Analysis

Few numbers are required for most case series studies.

Some may report percentages (as descriptions not outcomes) **such as:**

- **Case fatality rate** (is the proportion of persons with a particular disease who die as a result of that condition)
- **Mortality rate** (is the percentage of members of a population who die of any condition during a specified time period)
- **Proportionate mortality rate** (is the proportion of deceased (dead) members of a population whose death was attributable to a particular cause)

With sufficient sample size, comparisons can be made between subpopulations of cases.

CHAPTER 9: CROSS-SECTIONAL SURVEYS

9.1 Overview

A useful video

The most common study designs used

Cross-sectional survey = prevalence Study

Used in collecting data (one point in time), incidence (during a particular time period)

so it ≠

Measures the proportion of a population with a particular exposure disease.

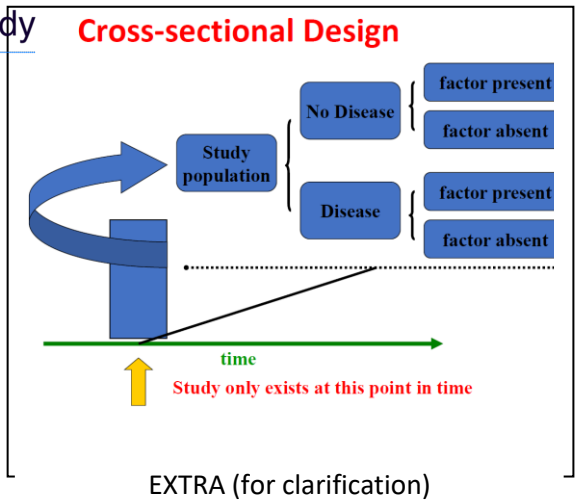
Or

Most popular approaches as they allow for collection of new data. (advantage)

Rapid

>> So, it's considered a study design, and a collection method.

Data



EXTRA (for clarification)

FIGURE 9-1 Key Characteristics of Cross-Sectional Surveys

Objective	Describe the exposure and/or disease status in a population
Primary study question	What is the prevalence of the exposure and/or disease in the population?
Population	The study participants must be representative of the population from which they were drawn.
When to use this approach	Time is limited and/or the budget is small.
Requirement	The exposures and outcomes are relatively common, and the researchers expect to be able to recruit several hundred participants.
First steps	1. Define a source population. 2. Develop a strategy for recruiting a representative sample. 3. Decide on the methods to be used for data collection.
What to watch out for	Non-representativeness of the study population
Key statistical measure	Prevalence

9.2 Representative Population

The participants must be reasonably representative of some larger population.

Example: If the results are intended to reflect the profile of an entire town, then the study's sampling strategy must recruit a population that's as diverse as the town.

>> So, the participants must be representative in terms of number and characteristics. It will be discussed later…

9.3 KAP Surveys

A KAP survey (a commonly used cross-sectional study type) **asks participants about their:**

Knowledge + Attitudes, beliefs, or perceptions + Practices or behaviors.

It can be helpful for identifying gaps between what people know and how they act on that knowledge.

9.4 Repeated Cross-Sectional Surveys

A repeated cross-sectional study re-samples & re-surveys representatives from the source population at two or more different time points.

This type of study doesn't track the same individuals forward in time. Rather, a new set of participants is sampled from the source population each time a survey is conducted. Some people may happen by chance to be selected for more than one round of surveying, but their answers to the different surveys aren't linked. So, it can reveal trends in population-level (not individual-level) metrics over time. يعني هذا النوع من الدراسات يقيس حجم تغيرات معينة في الناس بشكل عام وليس في كل شخص على حدة، ولو أردنا دراسة شخص واحد لاستخدمنا نوع أخر من الدراسات يسمى Longitudinal Cohort Study.

9.5 Analysis: Prevalence

We have previously discussed the prevalence which calculated at a one point in time, but sometimes we need to calculate the prevalence over a short duration of time, with all data collected within a few days, weeks, or months; therefore we call it the prevalence rate.

Prevalence Rate: the percentage of the population with a given trait at the time of the survey.

Prevalence rate ratio: ratios that compare prevalence of a characteristic in two population subgroups.

** Remember: causality (السببيَّة/Exposure --> Outcome) can't be established based on a cross-sectional study (because its function is only description of variables or population), but if we use correlational statistics, causality can be established (and this will be discussed later)…

Lecture 8

Previously , we discussed the first and the second types of the study designs ,today we will discuss the third one…..

3-Case-control study

Useful videos: 1

2

- A case-control study compares the exposure histories of people with and without a particular disease in order to identify likely risk factors for the disease
- .In this type of studies ,participants recruited based on disease status ,so we divide our participants into tow major groups:

1-**Cases**:participants who have the disease ,symptoms ,clinical manifestation

2-**Controls** :participants without the disease

- .Both (controls and cases)are asked the same set of questions about past exposures

Case-control studies are good for studying uncommon diseases (because these diseases need many years to occur so the researcher collect the data about the cases over years then he compares it with the controls)

FIGURE 10-1 Key Characteristics of Case-Control Studies

Objective	Compare exposure histories of people with a disease (cases) and people without that disease (controls)
Primary study question	Do cases and controls have different exposure histories?
Population	Cases and controls must be similar except for their disease status.
When to use this approach	The disease is relatively uncommon, but a source of cases is available.
Requirement	A source of cases is available.
First steps	1. Identify a source of cases. 2. Assign a case definition. 3. Decide what type of control population will be appropriate for the study. 4. Decide whether cases and controls will be matched.
What to watch out for	Recall bias
Key statistical measure	Odds ratio (OR)

11

- Notice that in case-control design we ask the same questions to two different groups (control and cases)

Steps to design a case-control study :

1-Finding cases & controls:

- All cases must have the same disease, disability, or other health-related condition as per the **case definition**
- Find cases through hospitals, specialty clinics, physicians' offices, public health agencies, disease registries, anddisease support groups.
- Use a **control definition** to ensure that controls are similar to the cases except for their disease status. (for example :if the cases are males between 20 and 40 years ,the controls should be the same)
- Find controls who are friends and relatives ofcases, hospital or clinic patients without the disease of interest, or members of the general population.

- **Matching** : how to match the cases and the controls, it has 3 types:

1-***No matching*** : in this type we don`t have matching criteria because these criteria participate in cofoundation of the association between the key exposure and the disease(these criteria will affect-as well as the key exposure- the accuracy of the results)
2- ***Frequency (group) matching:*** Select one or more controls per case who are similar by age, sex, or other characteristics, but do not match cases to particular controls. (many controls per 1 case)
3- ***Matched-pairs (individual) matching:*** Each case is personally linked to a particular individual control,(example: Recruit a genetic sibling or other control who islinked to a particular case during analysis.) (1 control per 1 case)

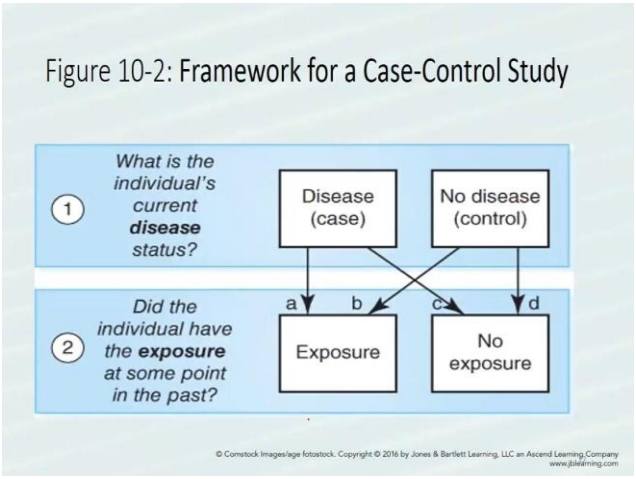
Avoid **overmatching** ,because it will be difficult to find controls who meet all the matching criteria (but if we do so we will end in a study population that is different from the general population ,so the EXTERNAL VALIDITY will be affected) → result in statistical bias that obscures the relation btw an exposure and the disease

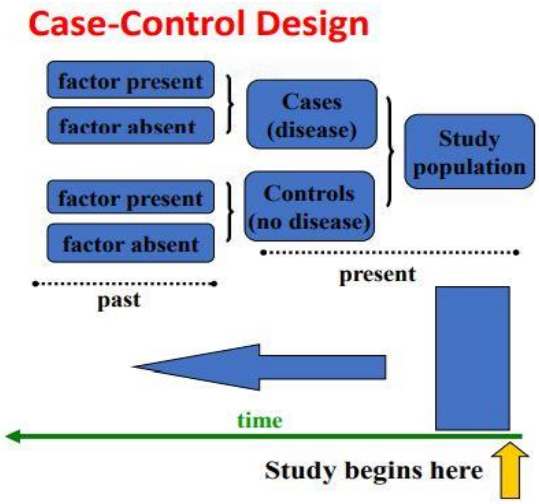
Special considerations

- Avoid ***misclassification bias*** with good case &control
- Be aware of ***recall bias,*** which occurs when cases & controls systematically have different memories ofthe past,(all the data that will be collected depends on the memory of the individuals)

2-Analysis :Odd ratios(ORs)

- As we said before, our population will be divided into two major groups(cases and controls)and Each one of these groups will be divided into another two groups (exposure and no exposure) ,so at The end we will end up with 4 groups :



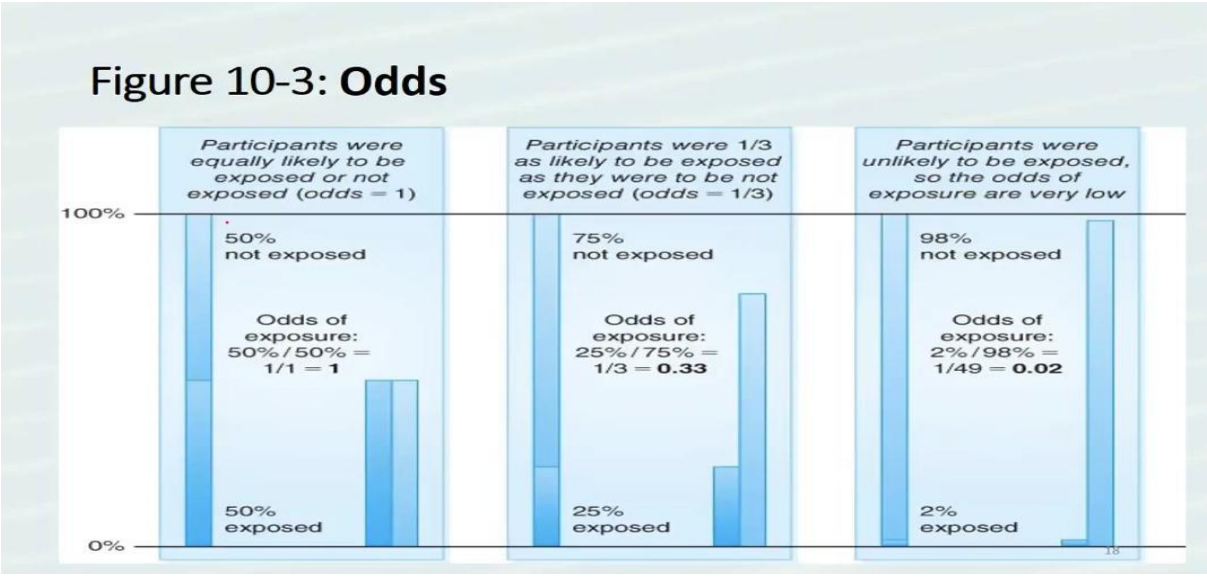


Now,

Odds: Compares the likelihood of having had a particular exposure to not having had it.

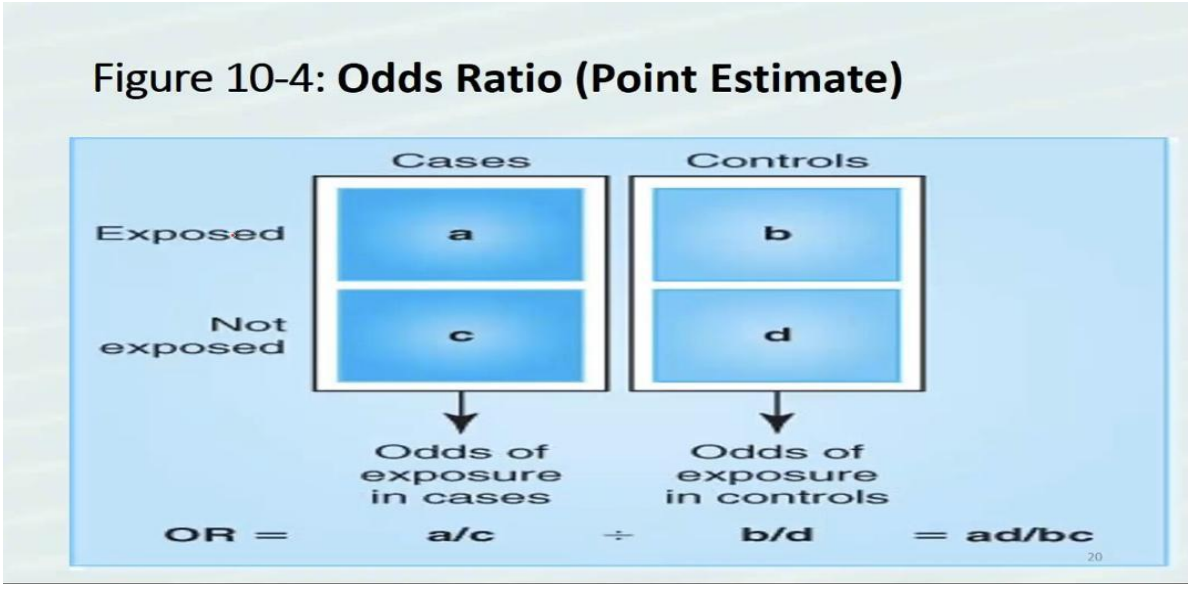
***Odds* ratio**: Compares the odds of exposure among cases to the odds among control The main measure of association in case-control study

- **The figure below can explain the meaning of "odds" :**



To compare two dichotomous (yes/ No) variables

- A **2*2** table displays the counts of people with various combinations of exposure status & disease status as follows:

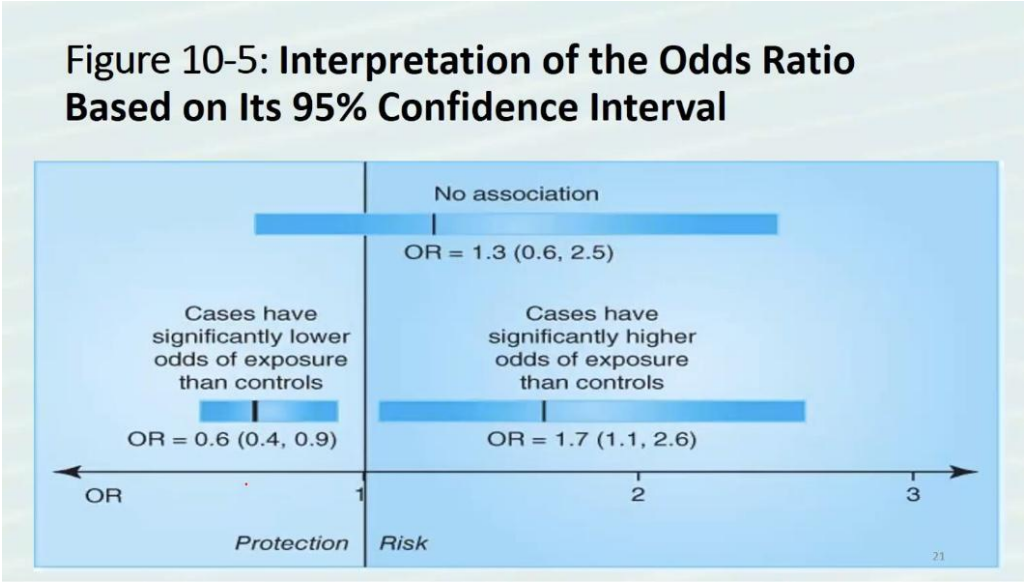


If OR:

=1 (the odds of exposures for cases and controls are the same)

>1 (Cases had higher odds of exposure than controls, implying that the **exposure was risky.)**

<1 (Cases had lower odds of exposure than controls, implying that the **exposure was protective.)**

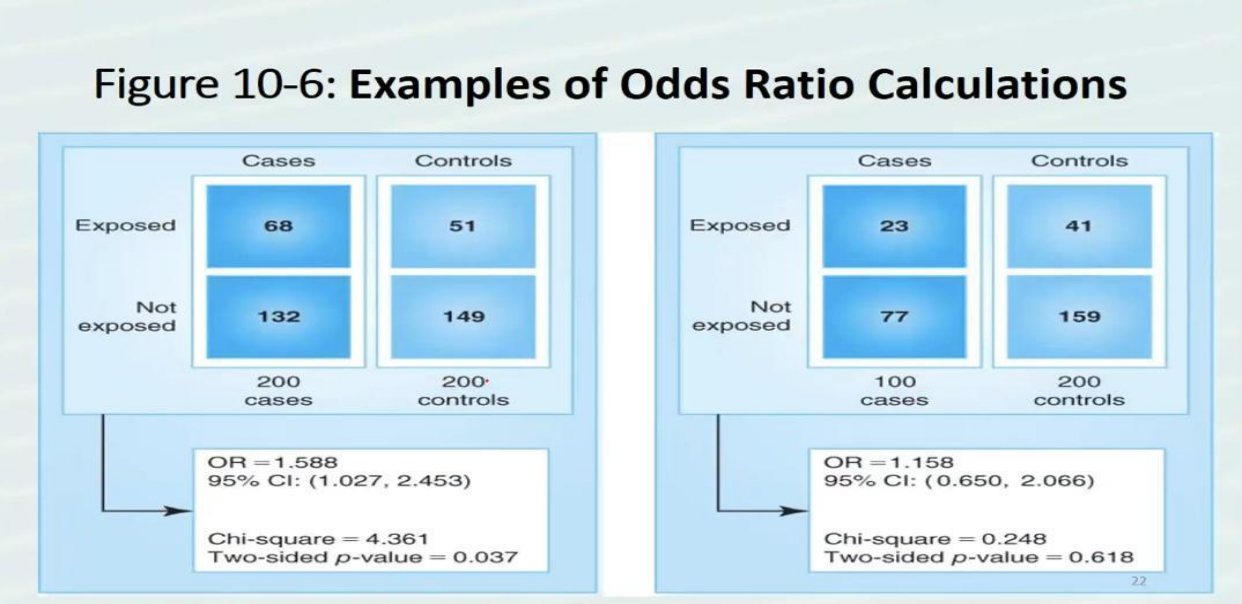


- Note : the following explanation is from the book, Dr. Jafar didn`t explain it 😊

IF the C.I is entirely lower than 1 -as the lower left one- ,then the odd ratio is statistically significant ,so the exposure is protective

IF the C.I is entirely more than 1 -as the lower right one- ,then the odd ratio is statistically significant ,so the exposure is risky

IF the C.I overlaps OR=1 -as the upper one- , then the odd ratio **not** statistically significant in study population



- We are not required to calculate Chi-square or p-value

P -value>0.05 indicates no association
P-value < 0.05 indicates statistically significant

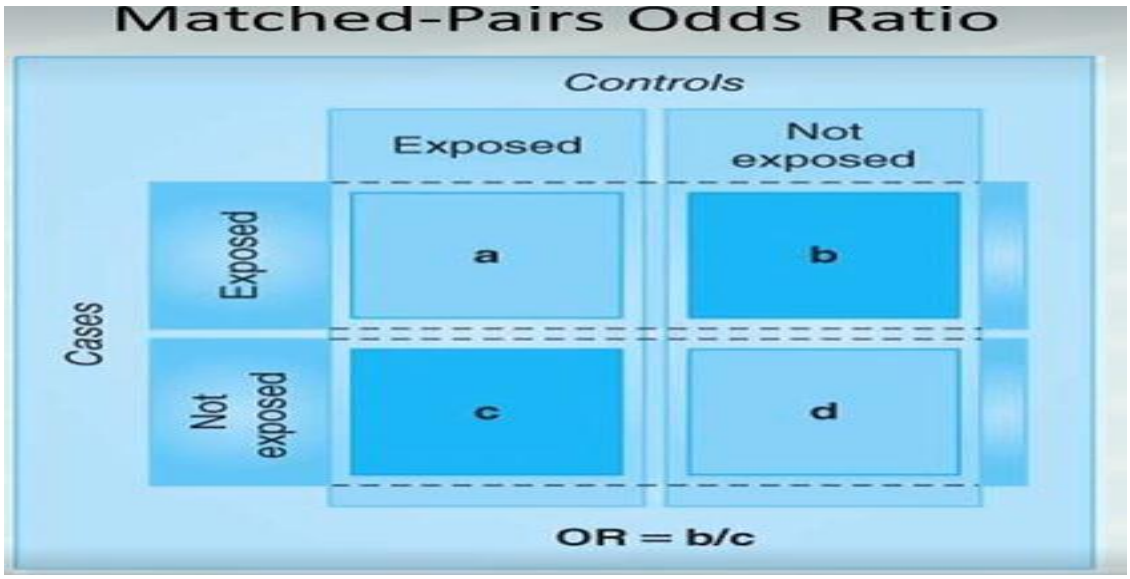
Matched case-control studies :

**It is a special type of case-control studies ,here we match -as much as we can – every control to every case(case by case should be matched)

**A special type of 2*2 table displays the distribution of pairs of cases and controls

- In **concordant** pairs :the case and control have the same exposure history
- In **discordant** pairs :the case and control have different exposure history

**NOTE: here we have more accurate and reliable results.



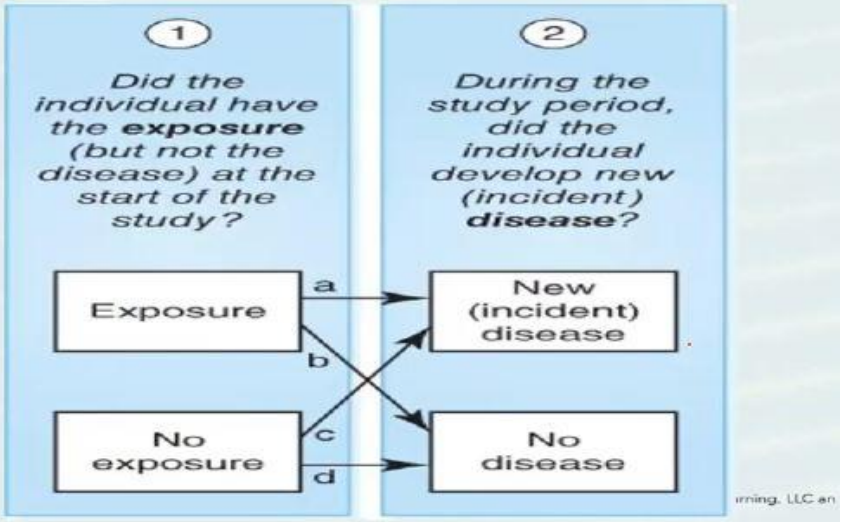
CHAPTER 11: COHORT STUDIES

11.1 Overview:

Useful videos: 1 2

- A cohort study follows participants through time to calculate the rate at which new disease occurs and to identify risk factors for that disease
- **Cohort** : is a group of similar people followed through time together
- **Cohort studies** are **observational** (not experimental) studies with at least two measurement times: a baseline and a follow-up examination
- Cohort studies quantify the **rate of incidence** (new) disease
- One of the most famous cohort studies is Framingham study

Figure 11-2: **Framework for a Cohort Study**



11.2 Types of Cohort studies:

We have **3** types of cohort studies :

1-**Retrospective(historical) cohort study** :recruits based on exposure status at some point in the past and uses follow-up data from some point after that old exposure to ascertain disease status

2-**Prospective cohort study** :recruits based on exposure status in the present and follows them forward the time

3-**Longtudinal cohort study** :recruits a representative sample of population and follows people forward in time(multiple exposures and multiple diseases)

Participants recruited based on member ship in a well - defined source population

FIGURE 11-1 Key Characteristics of Cohort Studies

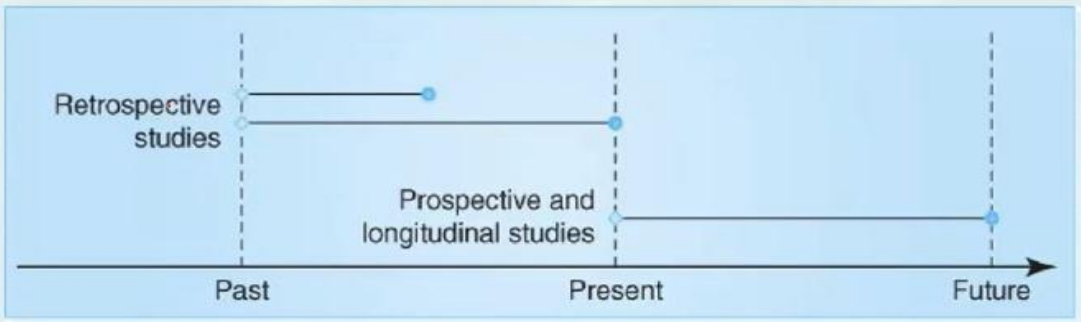
Approach	Prospective or Retrospective Cohort	Longitudinal Cohort
Objective	Compare rates of new (incident) disease over time in people with and without a particular well-defined exposure.	Follow a representative sample of a well-defined population forward in time to look for new (incident) diseases associated with a diversity of exposures.
Primary study question	Is exposure associated with an increased incidence of disease?	Is exposure associated with an increased incidence of disease?

FIGURE 11-1 Key Characteristics of Cohort Studies (continued)

Approach	Prospective or Retrospective Cohort	Longitudinal Cohort
Population	Participants must be similar except for exposure status.	Participants must be available for follow-up months or years after enrollment.
	Because the goal is to look for incident disease, no one can have the disease of interest at the start of the study.	The study participants must be reasonably representative of the population from which they were drawn.
Use this approach when	An exposure is relatively uncommon but a source of exposed individuals is available.	The goal is to examine multiple exposures and multiple outcomes and time is not a concern.
Do not use unless	A source of individuals with the exposure is available.	There is adequate time and money for the study.
First steps	1. Identify a source of individuals with the exposure. 2. Decide what type of unexposed individuals will be an appropriate comparison group.	1. Select a source population. 2. Select the exposures and outcomes that will be assessed. 3. Decide how often data will be collected. 4. Develop a strategy for minimizing the burden of participation and maximizing benefits and incentives.
Watch out for	Loss to follow-up (prospective studies) or missing records (retrospective studies)	Loss to follow-up
	Information bias in which the exposed participants are more thoroughly examined for disease than unexposed participants	Potential data management challenges if a lot of information is collected at many points in time
Key statistical measure	Incidence rate ratio (RR, also called the relative risk)	Incidence rate ratio (RR, also called the relative risk)

- There isn't a clear edge between the duration of prospective &longitudinal cohort studies, some books suggest that the duration of 6 months indicates prospective, after that we consider it longitudinal.

Figure 11-3: **Times of Baseline and Follow-Up Data Collection for Cohort Studies**



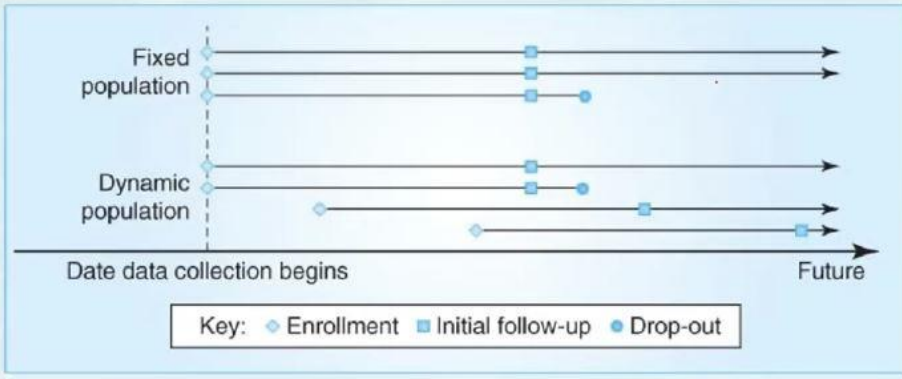
*Notice that in retrospective studies we divide the population into two groups (exposed and non-exposed -***in the past**-) and we follow up the two groups to find whether a certain disease will develop (or developed) or not

* Notice that in prospective studies we divide the population into two groups (exposed and non-exposed -***now**-) and we follow up the two groups to find whether a certain disease will develop or not

Also called time series studies or panel studies

- Longitudinal studies may use **a fixed population** or a **dynamic (open) population** with rolling enrolment

Figure 11-4: **Longitudinal Studies**



*IN fixed population ,all the population enroll the study at the same time (we start with all the participants)
*IN dynamic population ,the population enroll at different times (some of them with the beginning of the study ,some of them after one week ,some of them after 3 weeks and so on,,,)
*Notice that in both types of population the "Drop-out" may occur at any time
*If we have a study that aim to observe every participant in population for 3 years after enrolment the study ,notice that if the study uses dynamic population, we will end up with a study with a duration more than 3 years , because every new participant will have his own 3 years .

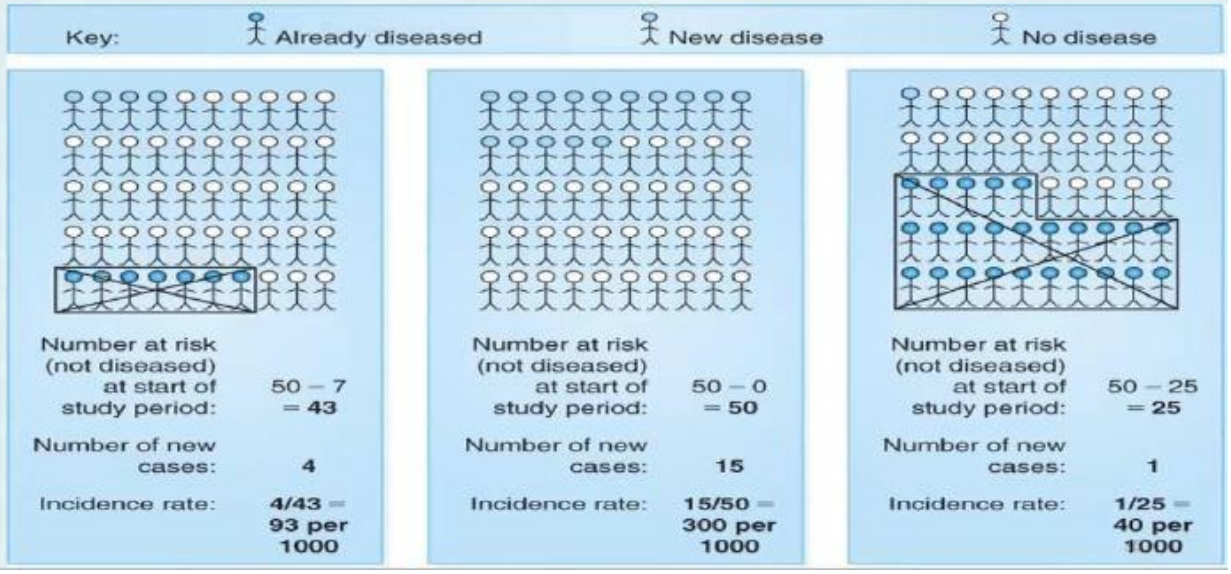
11.3 Special Considerations:

- Retrospective studies requires a source of valid data about past exposure status
- Prospective and longitudinal studies must take steps to minimize loss to follow-up when studies continue for many years

11.4 Analysis: Incidence Rate Ratios (RRs):

- The ***incidence rate:*** the number of new cases of disease in a population during a specified period of time divided by the total number of persons in the population who were at risk during that period.

Figure 11-5: Calculating *incidence rate*



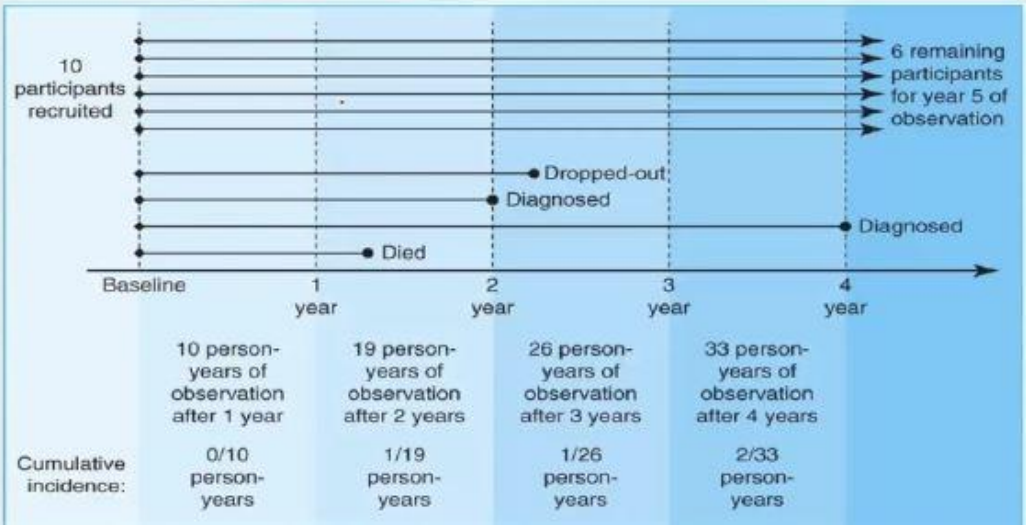
*I.R=no. of new cases * 1000 , notice that we exclude the already diseased participants from the
Population ***at risk*** population who are at risk

*we need to calculate I.R for exposed and unexposed groups

*For the studies with dynamic population we can use another method to calculate the incidence ratio which is called **person-time,** the explanation below is from the book ,,,,,

Some cohort studies, especially those with dynamic populations and those that run for many years, use person-time as a denominator. **Person-time** is a way of accounting for individuals in the study population being observed for different lengths of time. Person-time can be expressed in units of person-years, person-months, or even person-days. Suppose that a study recruits 10 individuals at baseline (**Figure 11-6**). After 4 years, 6 of the 10 participants are still active in the study and have not been diagnosed with the disease of interest. Together, these 6 individuals have contributed 24 person-years of observation during the first 4 calendar years of the study. Suppose that 2 of the 10 original participants are diagnosed with the disease of interest at their annual study examinations. One person is diagnosed 2 years into the study, and the other 4 years into the study. Together, these 2 individuals contributed 6 person-years of observation to the study. However, once they are diagnosed and no longer able to develop incident disease, they are no longer able to contribute further person-years to the denominator for the calculation of incidence. Two other participants also leave the study and are **censored** (removed from analysis). One drops out of the study after the second year but before the third year; this participant is considered to have contributed 2 person-years of observation. Another dies after the first year and contributes only that 1 person-year of observation. In total, over 4 calendar years, the 10 original participants experience 2 incident cases of disease over 33 person-years of observation. For the calculation of incidence rate ratios and other measures that rely on the comparison of incidence rates, it does not matter whether the incidence rates are measured per 1000 participants (Figure 11-5) or per 1000 person-years (Figure 11-6), as long as all incidence rates in the equation use the same units.

Figure 11-6: *Person-time*



**RR=Incidence rate among exposed
Incidence rate among unexposed

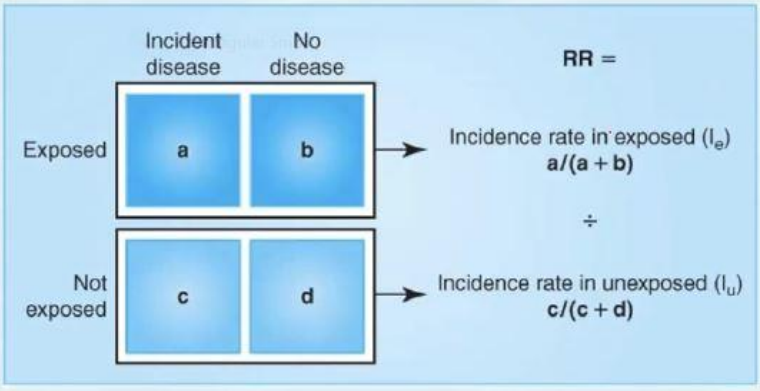
- **incidence rate ratio (RR):**

=1 The incidence rate was the same in exposed and unexposed groups

>1 The incidence rate was higher in exposed than unexposed ,indicating that the exposure was risky

<1 The incidence rate was lower in exposed than unexposed ,indicating that the exposure was protective

Figure 11-8: **Rate Ratio (Point Estimate)**



***We can calculate RR on its confidence interval (as same as OR from the previous lecture)**

Figure 11-9: **Interpretation of the Rate Ratio Based on Its 95% Confidence Interval**

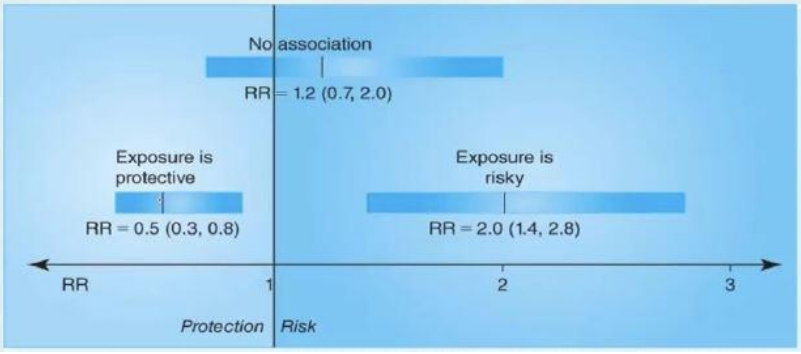
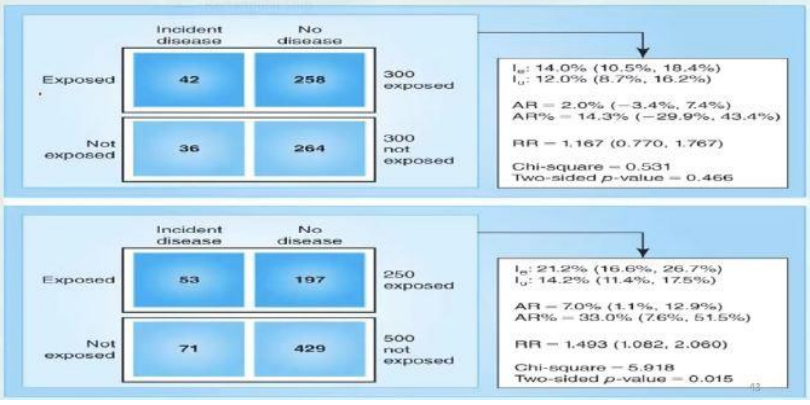


Figure 11-10: **Examples of Rate Ratio Calculations**



Note: we can calculate **Attributable (Excess risk) as follows :

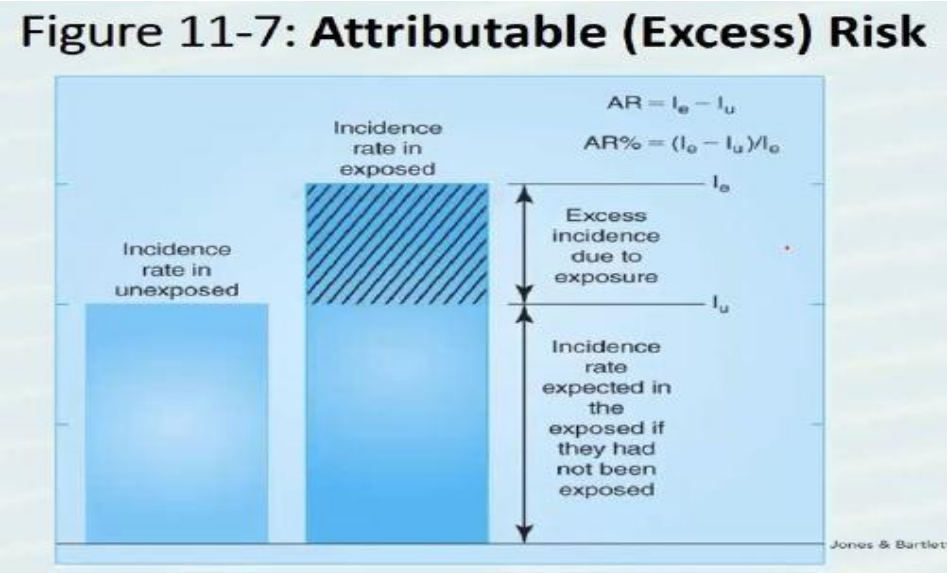
Attributable risk(AR)=Incidence in exposed-Incidence in nonexposed

Which is represented by the striped area below

Attributable Risk percent: is the proportion of incident cases among the exposed that are due to exposure ,or it is the proportion of the cases of the disease in the exposed that could have been prevented if the exposure was removed .

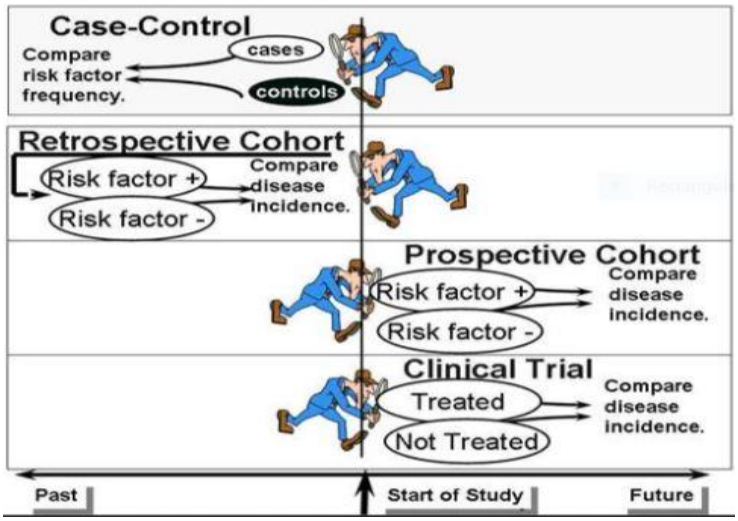
AR%= AR

Incidence in exposed



**To conclude : below, a picture shows the difference between case-control studies and cohort studies (Retrospective &Prospective), it isn`t from the slides :

A useful video to compare between case-control &cohort studies



"لا أعلم علمًا بعد الحلال والحرام أنبل من الطب إلا أن أهل الكتاب قد غلبونا عليه". الإمام الشافعي

اقرأ لترتقيَ بأمّتك