
MEDICAL RESEARCH
MODIFIED NO. 10
كتابة: لجين بلال وميس مصطفى
تدقيق: ميس مصطفى
الدكتور: منير أبو هلالة

Week 5 Descriptive studies part 2

Color code

- Slides
- Doctor
- Additional info
- Important

This part covers ecological studies and cross-sectional studies

Ecological studies

Are studies in which information on the characteristics and/or exposures of individual members of the population groups are generally not obtained. Existing statistics (from the World Bank or WHO) are used to compare the mortality or morbidity experience of one or more populations with some overall index exposure. Care is needed to avoid the 'ecological fallacy' where inappropriate conclusions are made from ecologic data

- Ecological studies are research approaches that focus on the characteristics and exposures of entire populations or communities rather than individual patients. In these studies, we often rely on large-scale data from organizations like the World Bank, WHO, and other agencies. For example, we may examine data on disease incidence and potential risk factors to explore correlations and identify possible risk factors. This allows researchers to generate hypotheses about what might contribute to certain health outcomes

Ecological studies

- These studies are used to describe disease or drug use problems in relation to some factor of interest.

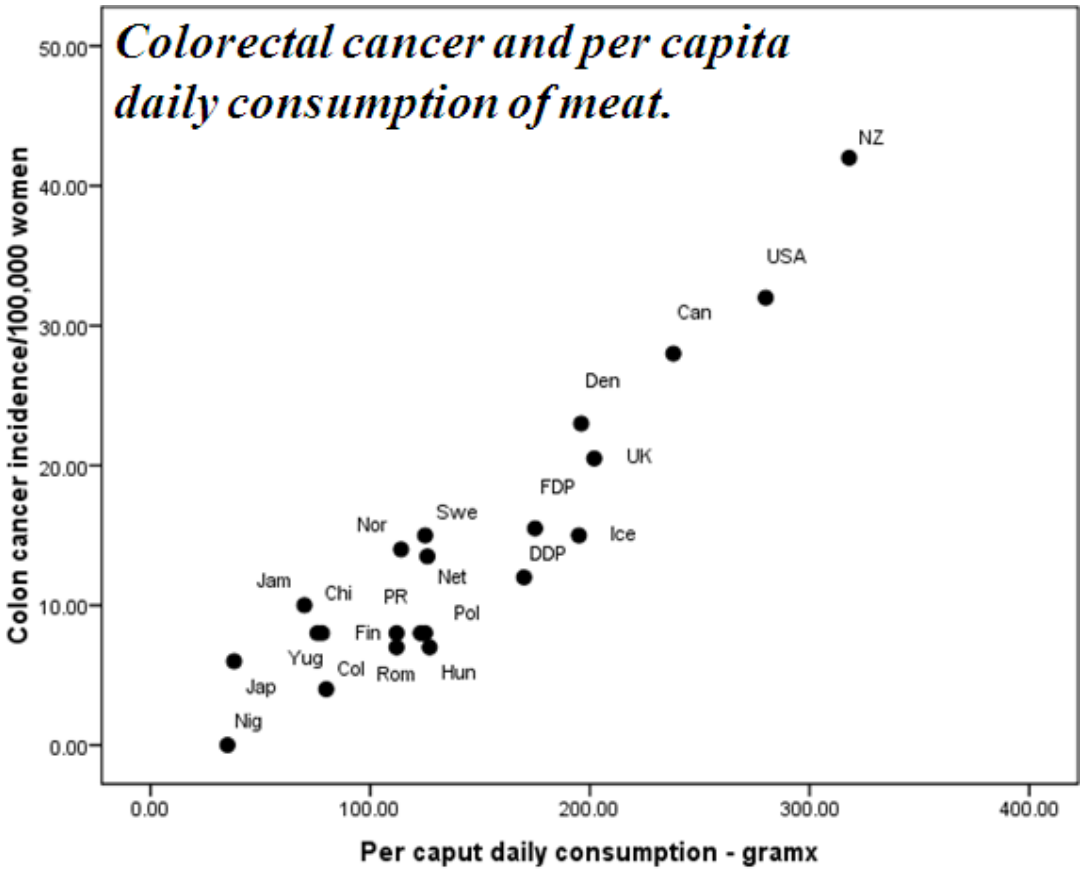
Comparing cigarette consumption with rates of cancer

Comparing Alcohol consumption with coronary heart disease mortality

- Ecological studies are the first identified strong relationships between disease and behavior.

Ecological studies

- In ecological studies the unit of analysis is some aggregate individuals rather than individual persons
- Geographic areas or time period are often used as a basis for defining aggregates
- The analysis centers on determining whether the ecological units with a high frequency of exposure are also unit with a high frequency of disease (+ve correlation) or a low frequency of disease (- ve correlation)

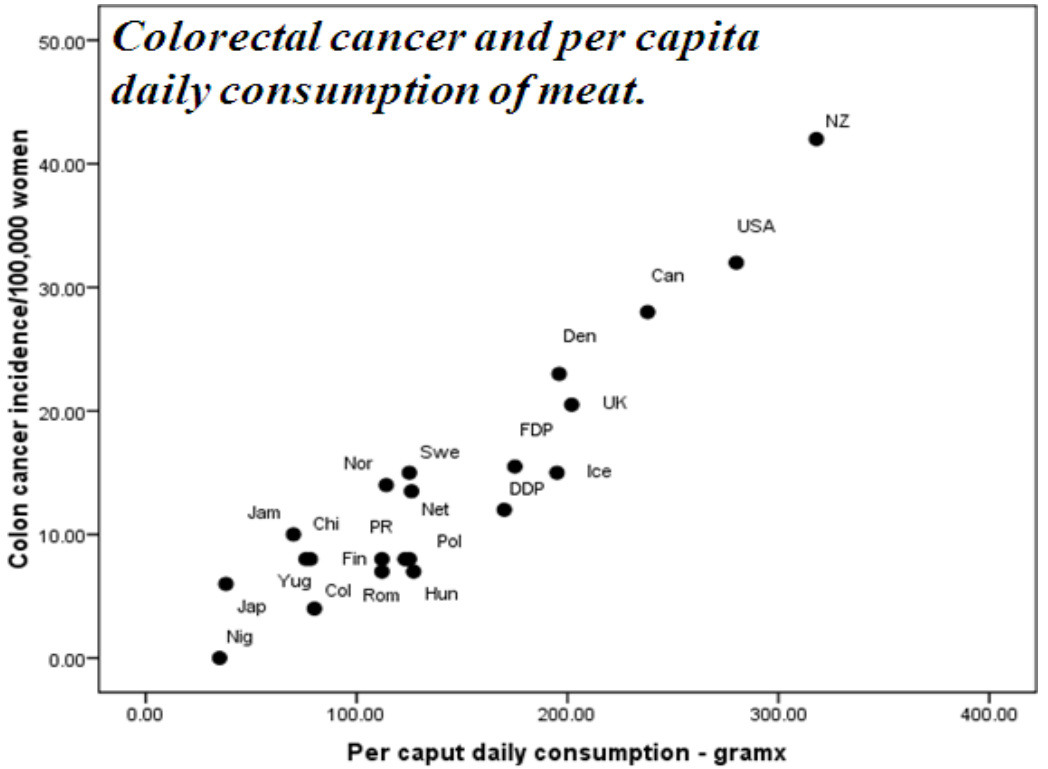


Adapted from: *Int. J. Cancer 15:617, 1973*

X-axis: red meat consumption
Y-axis: incidence of colon cancer

You can see from this graph that as we move along the x-axis (i.e. as red meat consumption increases), the higher the incidence of colorectal cancer. In countries where there is low meat consumption like Nigeria and Japan, there is a low incidence of cancer, and in countries where there is high meat consumption like the USA and New Zealand, there is a high incidence of cancer. This observation enables us to generate hypotheses and draw correlations that will be further tested using analytical studies.

- This is an example of a positive correlation



Adapted from: *Int. J. Cancer 15:617, 1973*

An example of a negative correlation would be:
x-axis: physical activity, y-axis: any disease (diabetes or ischemic heart disease, etc)
As we move along the x-axis (i.e. more physical activity), the lower the disease incidence. Thus, if we were to draw a line connecting the dots, it would be in the opposite direction than the one above.

These are great studies that are inexpensive, easy, and quick to complete. We have individuals known as data miners who can draw correlations, and when there is a possible risk factor, we must conduct additional research to prove or disprove this hypothesis. However, there are two main limitations:

First, they examine entire populations rather than focusing on individuals with specific diseases. For example, if we're studying red meat consumption in the U.S. or New Zealand, we might assess average consumption levels across the whole population at a specific point in time, without separating those who have colorectal cancer. This can be limiting because certain subgroups, like colorectal cancer patients, may have different consumption patterns. Additionally, factors like seasonal variations in diet and activity levels are not captured when we only look at a single time point.

The second limitation involves confounding factors—variables that might influence the outcomes we observe. For instance, in countries like the U.S. and New Zealand, low fiber diets, high obesity rates, and limited physical activity could all contribute to higher rates of colorectal cancer, rather than red meat consumption itself being the primary risk factor.
A classic example of confounding is seen in studies linking heavy alcohol intake with lung cancer. Initially, heavy alcohol drinkers appeared to have a higher risk of lung cancer compared to non-drinkers. However, once researchers accounted for smoking, they saw that among heavy drinkers, only those who also smoked had a significantly increased risk of lung cancer. This showed that smoking—not heavy drinking—was the true risk factor.

In summary, the two main limitations of ecological studies are that they 1) focus on whole populations instead of disease-specific individuals and 2) are susceptible to confounding factors. Nonetheless, ecological studies are valuable for generating hypotheses, which can be further tested in more controlled clinical studies.

Ecological (correlational studies)

- look for associations between exposures and outcomes in populations rather than in individuals.
- They use data that has already been collected.

- The measure of association between exposure and outcome is the correlation coefficent *r*.
- This is a measure of how linear the relationship is between the exposure and outcome variables. (Note that correational is a specific form of association and requires two continuous variables)

To characterize illnesses, drug issues, and various correlations, these studies examine the relationship between exposure and results in a population rather than in individuals and at a certain point in time

Ecological (correlational studies)

Advantages of an ecological study

1. An ecological study is quick and cheap to conduct.
2. It can generate new hypotheses.
3. It can identify new risk factors.

Ecological studies have been instrumental in identifying various socioeconomic and occupational cancer patterns. While these studies don't provide definitive evidence to confirm or rule out specific risk factors, they help generate ideas and identify potential links. These preliminary findings can then be tested more rigorously through analytical studies, like case-control and cohort studies, to validate or refute the hypotheses.

Ecological (Correlational studies)

Disadvantages:

1. It is unable to control for confounding factors. This is often referred to as 'ecological fallacy', where two variables seem to be correlated but their relationship is in fact affected by cofounding factor(s).
2. It cannot link exposure with disease in individuals as those with disease may not be expose.
3. Its use of average exposure levels masks more complicated relationships with disease.
4. Its units of study are populations not individuals. Therefore, the disease rates linked with population characteristics and the association observed at group level does not reflect association at individual level.

Ecological (correlational studies)

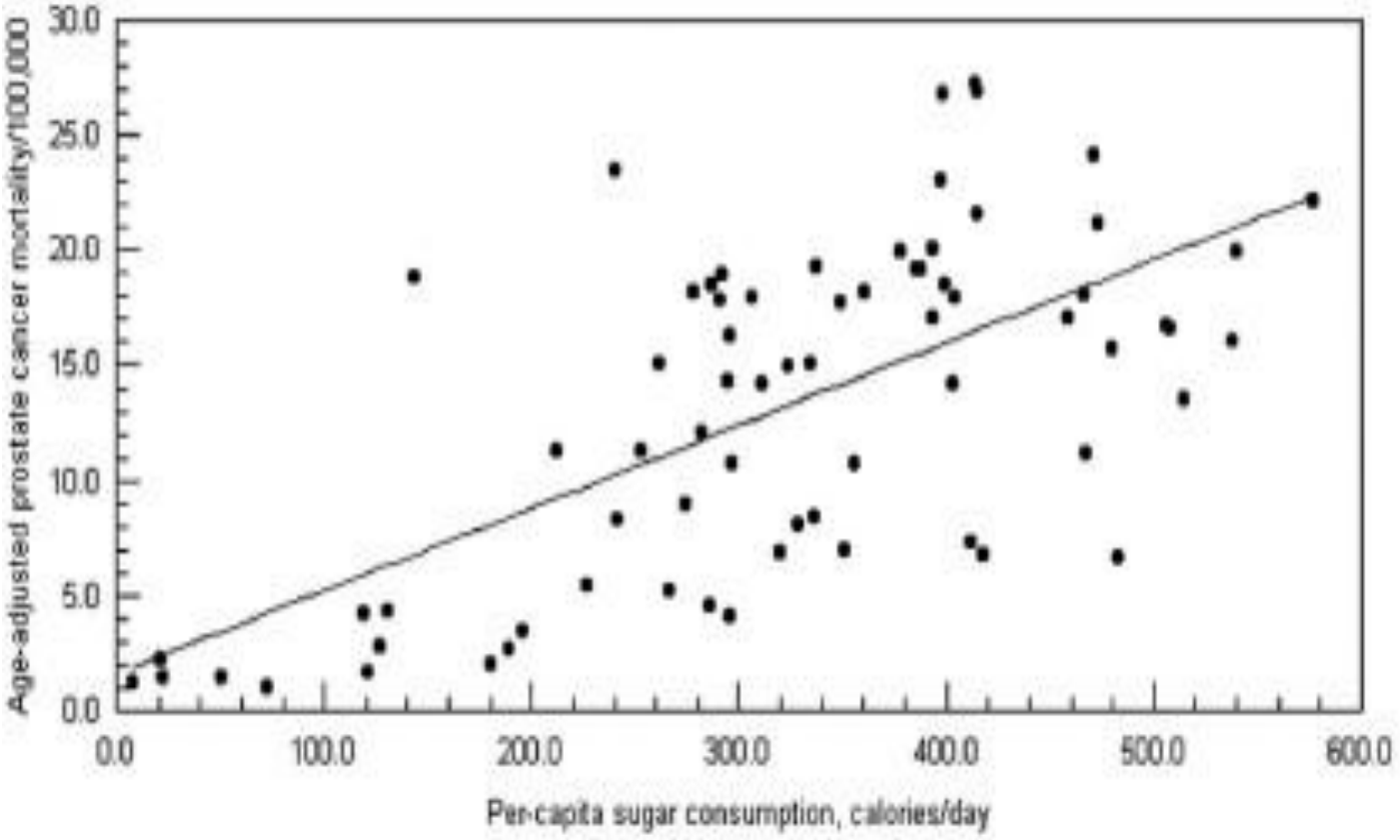


Fig. 1. Prostate cancer mortality versus sugar consumption in 71 countries.

Countries with lower sugar intake have lower mortality rates due to prostate cancer

Descriptive epidemiology

- There are many problems with descriptive methods.
- In case reports and case series, there is no control group.
- For correlation studies: there are confounding factors that might mask the true impact of risk factors.
- Correlation studies present only a snapshot of the problem, such as disease or drug use, in a population.

In descriptive epidemiology, case reports and case series lack control groups, and correlation studies, like ecological studies, often face issues with confounding factors. To establish stronger evidence, analytical studies are needed to test hypotheses. For example, Professor Howsen observed that 17 out of 20 cervical cancer cases in a case series tested positive for HPV. He then conducted a cross-sectional study comparing HPV prevalence in women with cervical cancer to women hospitalized for other reasons and women from the general population. His findings showed that HPV positivity was 80% among cervical cancer patients but only 20% in the general population and among those with other conditions. Similarly, if we're studying the correlation between smoking and hypothyroidism, an analytical study might reveal that smoking prevalence is 70% in hypothyroid patients versus 20% in the general population, suggesting smoking could be a risk factor. Conversely, if smoking prevalence were the same (e.g., 50%) in both groups, further investigation would likely be unnecessary.

CROSS-SECTIONAL STUDY DESIGN

- **Sometimes called *prevalence studies*.**
- **They are studies of total populations or population groups in which information is collected about the present and past characteristics, behaviors, or experiences of individuals.**
- **There are a number of advantages in performing a cross-sectional study.**
- **These studies involve a single data collection and, thus, are less expensive and more expedient to conduct.**

Cross-sectional studies are particularly valuable for assessing the burden of illnesses within a population. For example, if we want to understand the prevalence and impact of chronic conditions like type 2 diabetes, ischemic heart disease, or hypertension in Jordan, cross-sectional studies provide insight into the extent of these issues. Consider a cross-sectional study conducted to measure the prevalence of hypothyroidism in Jordan, specifically in areas like Karak, Irbid, and Amman, with 10,000 participants. Initially, each participant was asked if they had hypothyroidism, identifying 1,000 known cases. Then, TSH and T4 screenings were performed, revealing an additional 500 cases. To calculate prevalence, both known and newly identified cases (1,000 + 500 = 1,500) were divided by the total sample (10,000), resulting in a prevalence rate of 15%.
Cross-sectional studies like these are essential for estimating disease prevalence, identifying risk factors, and evaluating complications, making them powerful tools for understanding public health burdens.

Cross-sectional (or prevalence) studies

Are studies in which a defined population is surveyed and their disease or exposure status determined at one point in time

- **The prevalence rates of disease in the whole population as well as in those with and without the exposure under investigation can be determined**
- **Cross-sectional studies are generally not suitable for a disease which is rare or of short duration as few people will have the disease at any one point in time**

To improve a country's healthcare sector effectively, understanding the distribution and prevalence of various illnesses is essential. Each country has unique health challenges, and Jordan's major health issues differ from those in neighboring regions and globally. Setting healthcare priorities requires knowing the prevalence of different conditions.

If we study diabetes prevalence, for example, with a sample of 1,000 subjects, we might find 100 known cases and identify 50 or 100 new cases, giving us a 20% prevalence. Cross-sectional studies are useful for estimating the burden of common, long-lasting diseases because they allow us to assess prevalence at a specific time. However, they are unsuitable for rare diseases due to the large sample sizes required. For instance, with a rare disease that occurs in 1 out of 100,000 people, we would need a sample of 10 million people to identify 100 cases. For these diseases, we instead use cohort studies to calculate incidence, as in the case of congenital heart disease, where we might track pregnant women over two years to determine the number of cases and then calculate incidence.

For risk factors, the type of study depends on the disease's prevalence and duration. To study the risk factors for rare diseases, case-control studies are more efficient. Conversely, for common diseases, cohort studies are better suited. For chronic or long-term conditions, cross-sectional studies help measure prevalence, while cohort studies can provide insights into incidence for rare diseases. In medical textbooks, for example, diseases like ischemic heart disease and diabetes are discussed in terms of prevalence, while rare diseases such as epilepsy or Parkinson's are described using incidence because they have lower occurrence rates.

For acute, short-duration diseases like the flu or fractures, prevalence studies are often ineffective, as point or period prevalence at a given time may show zero cases, giving an inaccurate picture of the disease burden. To assess the burden of these short-duration conditions, we instead calculate incidence. For example, to evaluate flu cases among medical students, we could track flu cases from November to April, confirm diagnoses with PCR tests, and calculate the incidence rate.

Ultimately, when studying the burden of a disease, acute or rare conditions require cohort studies to assess incidence, while long-term or common diseases benefit from cross-sectional studies to gauge prevalence. However, self-reported prevalence may underrepresent the true burden, as many patients are unaware of conditions like type 2 diabetes or hypertension until they reach more advanced stages. Thus, an accurate prevalence assessment requires combining self-reports with diagnostic testing, especially for conditions with high rates of undiagnosed cases.

In summary:
Burden of rare and acute diseases =cohort
Burden of common diseases =cross-sectional
Risk factors of a rare disease = case-control
Rare risk factors =cohort
Burden of acute infections/short-lasting illnesses and rare diseases = Incidence
Burden of common and long-lasting/ chronic illnesses = Prevalence

CROSS-SECTIONAL STUDY DESIGN

- **Emphasis is on differences between groups at one point in time.**
- **They provide a one-time glimpse at the study population, showing the relative distribution of conditions, diseases, and injuries—and their attributes—in a group or population.**
- **Point prevalence versus Period prevalence**

Point prevalence refers to measuring the number of cases at a specific point in time. For example, tomorrow, we could screen 10,000 people across northern, central, and southern Jordan for hypothyroidism, diabetes, ischemic heart disease, and hypertension.

Period prevalence, on the other hand, involves tracking the number of cases over a set period—say, over the next three months—to capture how many people have the condition during that timeframe.

Incidence is critical for understanding disease over time. For instance, if we find 100 cases per 10,000 people over six months, we would estimate an annual incidence of 200 cases per 10,000 by doubling the six-month rate. Conversely, if 100 cases are found over two years, the annual incidence would be 50 cases per 10,000.

In summary, prevalence counts both old and new cases in a population at a given time without focusing on a specific duration, whereas incidence considers the rate of new cases over a defined period.

Cross-sectional studies

- More effective in identifying chronic diseases and problems
- Less effective in identifying communicable diseases of short incubation periods and short durations.

Cross-sectional (or prevalence) studies

- **It is often difficult to separate cause and effect as the measurement of exposure and disease at any one point in time**

- **Because of this limitation, cross-sectional studies are useful when investigating exposures which do not change e.g genetic characteristics such as ABO blood group and HLA**

- **Cross-sectional studies are often used as an initial exploration of a hypothesis prior to conducting a case-control or follow-up study**

Understanding disease pathophysiology highlights that certain conditions, like ischemic heart disease, can begin developing long before symptoms emerge. For example, recent reports suggest that late teenagers may already start the early stages of ischemic heart disease or thrombosis. Similarly, some cancers may begin developing undetected, making it difficult to pinpoint when the disease process or a related risk factor, like smoking, actually began.

In a cross-sectional study of lung cancer patients, for instance, it's unclear whether smoking preceded the onset of cancer or if both developed concurrently. Cross-sectional studies lack the ability to establish a "temporal relationship" because they don't track exposures and outcomes over time. In contrast, cohort studies—which we will cover next week—allow for this temporal assessment. In cohort studies, participants are categorized at baseline based on their exposure (e.g., smokers vs. non-smokers) with those already having ischemic disease or diabetes excluded to ensure the disease is not already present. After establishing baseline health, we then follow the participants over time to observe which exposures precede disease development, providing a clearer temporal relationship.

So, in cross-sectional studies, it's challenging to determine the sequence of exposure and outcome—something cohort studies are specifically designed to address.

- Typically, we don’t rely on cross-sectional studies to establish risk factor relationships or determine temporal relationships. However, there are exceptions, such as blood type (A, B, AB, O) and certain genetic factors. Since we know blood type is present from birth, we can be confident that blood group O, for example, existed before the onset of conditions like peptic ulcers. Similarly, BRCA1/2 mutations, which are inherited, allow us to observe a high prevalence in breast cancer patients compared to non-patients, supporting a temporal relationship. Because these genetic factors and blood types are present from birth, we know they precede disease development, allowing us to use cross-sectional studies to assess chronic disease risk and prevalence. However, for short-duration infectious diseases or rare conditions, cohort studies are preferred to accurately gauge disease burden and impact.

- In cross-sectional studies, it's crucial to have a sample that represents the general population accurately. For example, if I conduct a study in a small village where the average age is 55-65 and find that 40% of residents have type 2 diabetes, it wouldn't be appropriate to conclude that the prevalence of type 2 diabetes in Jordan is 40%. This village sample skews older, so it doesn’t reflect the broader population. To represent the country accurately, a study should include samples from across regions—north, middle, and south—as well as from various areas within cities, diverse socioeconomic groups, and both rural and urban communities.

CROSS-SECTIONAL STUDY DESIGN

- **They provide information and data useful for the planning of health services and medical programs.**

- **Assessment of the burden of diseases or healthcare programs leads to setting priorities at the organization, local or national levels.**

- **They are based on a sample of the whole population and do not rely on individuals presenting themselves for medical treatment**

- They are useful for generating hypotheses. If we observe that certain factors are more prevalent among patients than in the general population, it can guide hypothesis generation. By examining the distribution of various risk factors, we can identify which ones to target in future programs. For example, since smoking rates are notably high in Jordan, we would know to implement programs aimed at reducing smoking rates.

CROSS-SECTIONAL STUDY DESIGN

- Sample size:

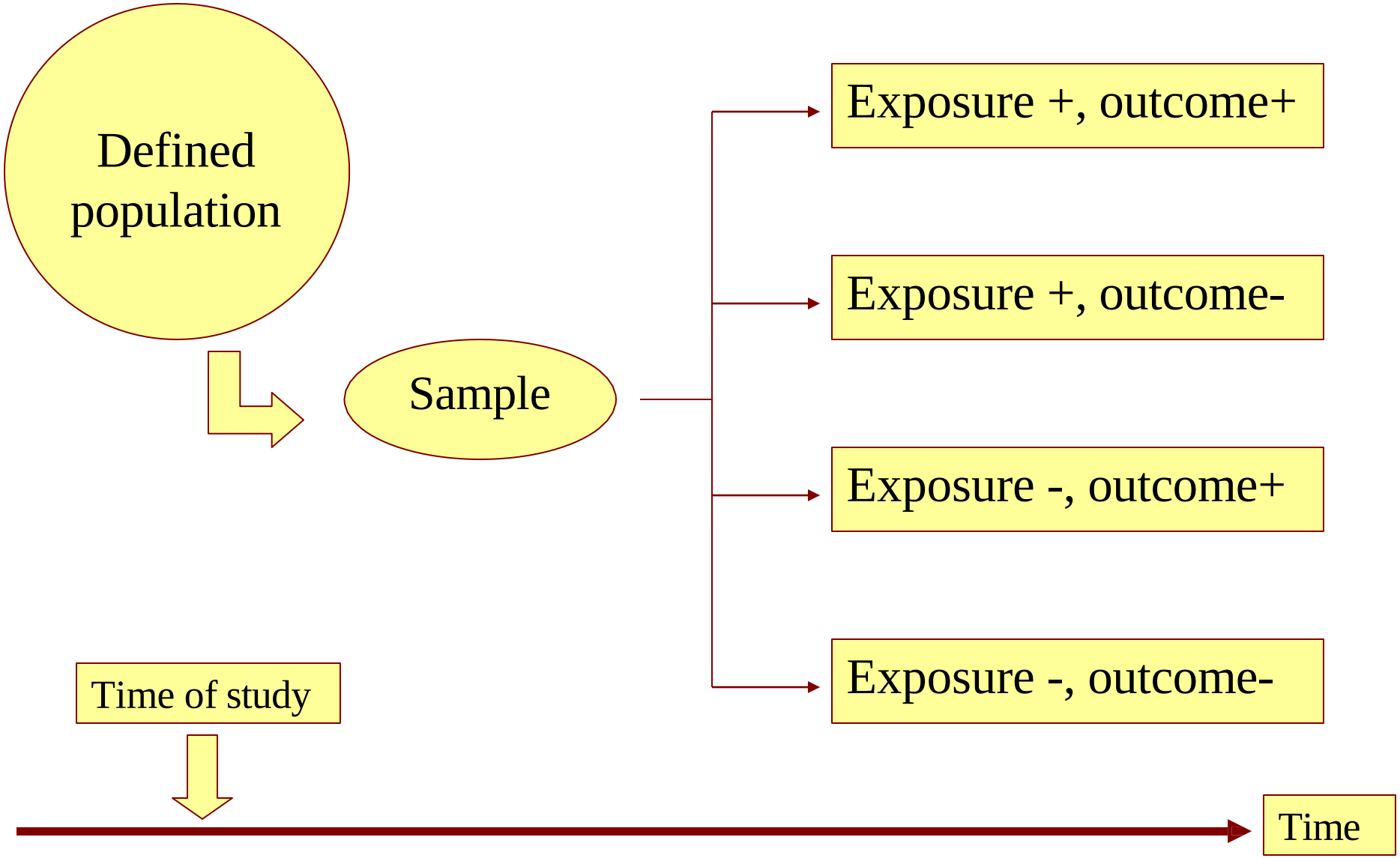
1. Question or primary & secondary outcomes
2. Population size
3. Prevalence of condition of interest in the population
4. Distribution of the condition (for example hypothyroidism is common among women age 50 to 70 but less common amongst men at this age group).

   Therefore we need a large sample from men in the general population to get men with hypothyroidism. In this case we stratify for gender.

Cross-sectional study

- Exposure and outcome are assessed simultaneously among individuals in a defined population, thus at one point in time
- No sampling of individuals based on a exposure or an outcome

Cross-sectional study



Two by two table

Exposure	**Outcome**		Total
	Yes	No	
Yes	a	b	a + b
No	c	d	c + d
Total	a + c	b + d	a + b + c + d

Prevalence of outcome in exposed = a / a + b

Prevalence of outcome in non-exposed = c / c + d

Prevalence Rate Ratio (PRR) = $= \frac{a / a + b}{c / c + d}$

Cross-sectional study

Prevalence of and Factors Associated With Persistent Pain Following Breast Cancer Surgery

JAMA. 2009;302(18):1985-1992

Objective To examine prevalence of and factors associated with persistent pain after surgical treatment for breast cancer.

Design, Setting, and Patients A nationwide cross-sectional questionnaire study of 3754 women aged 18 to 70 years who received surgery and adjuvant therapy (if indicated) for primary breast cancer in Denmark between January 1, 2005, and December 31, 2006. A study questionnaire was sent to the women between January and April 2008.

Cross-sectional study

Chemotherapy	Outcome		Total
	With pain	Without pain	
Yes	664	556	1220
No	879	1088	1967
Total	1543	1644	3187

Prevalence of pain among chemotherapy = 664/ 1220 = 54.4%

Prevalence of pain among no chemotherapy = 879 / 1967 = 44.7%

Prevalence Rate Ratio (PRR) = = 54.4 / 44.7 = 1.22

Cross-sectional survey of CHD among male by physical activity

	Number examined	Number with CHD	prevalence
Not physically active	**89**	**14**	**157.2/1000**
Physically active	**90**	**3**	**33.3/1000**

From: *BRCA1* and *BRCA2* genes mutations among 200 high risk breast cancer patients in Jordan

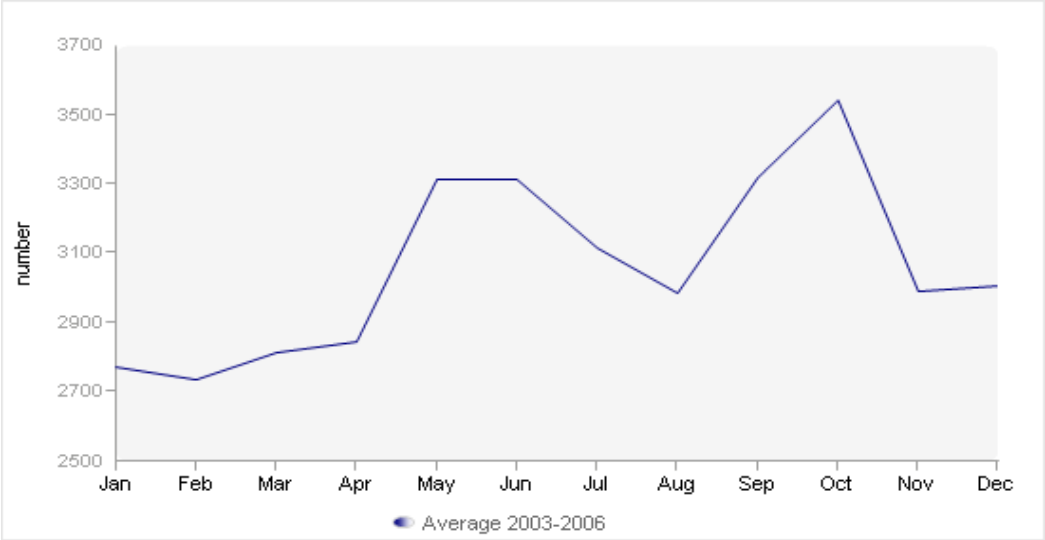
Category	Number of patients	Prevalence (total 200)
Recurrent mutations		
BRCA1 Positive	15	7.50%
BRCA2 Positive	14	7.00%
BRCA1 or BRCA2 Positive	29	14.50%
Possible (recurrent and novel) mutations		
BRCA1 Positive	7	3.50%
BRCA2 Positive	14	7.00%
BRCA1 or BRCA2 Positive	21	10.50%
Recurrent and novel (VUS and pathogenic) mutations		
BRCA1 Positive	15	7.50%
BRCA2 Positive	21	10.50%
BRCA1 or BRCA2 Positive	36	18.00%

Abu-Helalah et al. https://www.nature.com/articles/s41598-020-74250-2

Cross-sectional studies

- Seasonal variations of disease are not well represented in cross-sectional studies except if the duration of the study allows such comparison
- In the example below, studying RTA in October would not provide a valid result for incidence of RTA in whole year and does not allow identifying seasonal variations in the RTA
- Road traffic accidents by month of accident, Slovenia, average 2003-2006

Cross-sectional studies, unlike cohort studies, don't typically allow for the assessment of seasonal variations, which is a limitation. To study seasonal trends—such as for influenza—data collection over a full year is necessary. If a study is conducted only from December to March, it wouldn't capture seasonal variation throughout the entire year. Therefore, to assess seasonal changes, it's preferable to conduct a cohort study over at least one year rather than relying on a cross-sectional approach.



Cross-sectional studies: advantages

- **Relatively quick**
- **Data on all variables is only collected once.**
- **Sample size depends on the question**
- **Standard measures used**
- **Prevalence estimated**
- **The prevalence of disease or other health related characteristics are important in public health for assessing the burden of disease in a specified population and in planning and allocating health resources.**
- **Good for descriptive analyses and for generating hypotheses**

Cross-sectional studies

Disadvantages:

- They cannot show cause–effect relationships.

Difficult to determine whether the outcome followed exposure in time or exposure resulted from the outcome.

- If the sample is not representative, results are representative only of the individuals who participate in the study

Example prevalence of sickle cell anaemia in the Easter region of the KSA does not represent the who country.

- Not suitable for studying rare diseases or diseases with a short duration.
- Unable to measure incidence
- Associations identified may be difficult to interpret.
- Susceptible to bias due to low response and misclassification

To understand how to calculate prevalence or conduct cross-sectional studies, it's essential to examine research on the prevalence of various illnesses in Jordan and globally. This approach helps us analyze the distribution of illnesses, identify risk factors, and assess complication rates. For example, effectively managing complications of type 2 diabetes in Jordan requires an understanding of the prevalence and distribution of these complications, identifying specific risk factors, and determining if certain groups are more affected by these complications.

Additional sources

1. Book pages
2. Youtube videos
3. Webpages…etc

أستغفر الله العظيم وأتوب اليه

VERSIONS	SLIDE #	BEFORE CORRECTION	AFTER CORRECTION
V1→ V2			
V2→V3			

امسح الرمز و شاركنا بأفكارك لتحسين أدائنا !!