Chapter 1: The Health Research Process

1.1 The research process:

Research is the process of **systemically** and **carefully** investigating a subject in order to discover new insights about the world.

The research process is composed of 5 steps:

1. Identify a research question
2. Select a general study approach
3. Design the study and collect data
4. Analyze data
5. Write and share a report about the findings

- no matter what the goals of a research project are or what methods are used to achieve those goals, the five steps of the research process are the same.
- the first two steps are often completed concurrently (في نفس الوقت)

1.2 Health Research:

- **Health research** examines a broad spectrum of biological, socioeconomic, environmental, and other factors that contribute to the presence or absence of physical, mental, and social health and well-being.
- **Population health research** involves **humans** as the unit of investigation, rather than focusing on molecules, genes, cells, or other smaller biological components.
- Population health research ranges from clinical case studies with just a few individuals to global public health studies that may include many thousands of participants.
- Includes many sciences: demography, epidemiology, sociology, immunology etc.

- some studies that are very specific to one population at one place and in one point in time are not particularly helpful for identifying broader patterns. However, most health researchers hope that their findings will reveal trend, relationships, and theories that are generalizable to other populations, places, and times.

↳There is a distinction between routine practice activities and health research, sometimes routine acts in the hospital look as if they are research when they're not

Examples: An outbreak of gastroenteritis took place in a hospital, and they started looking into causes, is this a medical research? The answer is No, because as we previously said, a research is a question and a systematic way of finding the answer. Not just any routine investigation. *Usually such outbreak is not considered a scientific phenomenon it's just a mistake made somewhere*

Another example: satisfaction surveys of a hospital's service quality, which ask the patients how happy they are with the service given, but this is just a hospital feedback query to better their work not a research survey. However, if a group of researchers theorized an intervention that would make the service better and studied it, then it is considered a research.

1.3 Health Research Purposes:

- Needs assessment (community health profiles): What is the health status of this population? What are the major health concerns of members of this population? What health-related needs in this population are not being addressed?
- Risk assessment (risk factors for disease): What are the threats to health in this population? What are the risk factors for morbidity (illness), mortality (death), disability, and other health issues?
- Applied practice (clinical effectiveness): How well are we preventing, diagnosing, and treating health concerns in the populations we serve?
- Outcomes evaluation (impact of interventions): Of procedures, acts, projects, or educational programs used on this population. Is it effective? Is it not?

1.4 Book overview:

- Everyone can do meaningful research!
- The best way to learn about health research is to do real research.

Chapter 2: Selecting A General Topic

2.1 Practical Questions:

- Questions derived from **clinical practice**, community observations, and personal experience often point toward an unmet demand for needs assessments, program evaluations, and clinical effectiveness studies.

➔ Your environment, either you're a student in the classroom, a member in the community or whatever, from a self-point of view of interest as an individual, you might come up with research ideas.

- A good research question:

  1. **ends in a question mark:**

  ↳ This point is arguable, as there's different ways of reporting research in general. When you right a proposal for people to assess and look at to give you permission to do it or give you money, the way you write it is different from the way you write it when you report it as a final result in a journal for example...

  ↳ (Research problem/Research purpose/Research question) They refer interchangeably to each other.

  The difference is the way you state it only:

- A Question ends up with (?) and comes up as a question (what is? Is there? Etc...)

- A Purpose is usually prospective and carries a present tense in a way that investigates a certain clinical or basic problem. It addresses disease, exposure, population and sometimes settings and time frame of doing the research.

2. **Is testable:**

   Can be measured and examined – no measurement tool means no real findings.

2.2 Brainstorming and Concept Mapping:

- Use **brainstorming** to create a long list of possible research topics.
  People who are in the same field look at the problems that are usually encountered in their field, and they decide which project they're doing.
  This is not the stage for eliminating ideas because they do not appear feasible, and the ideas do not need to be well formed.

- Use **concept mapping** to identify central themes that might be worth exploring.
  It's complementary to brainstorming, they usually put the concept into its components manner. So, they put the major concept for example anemia, then we branch the anemia into different types, and the one type branches to many subtypes of the main type and causes,

Concept mapping: Narrowing down the main research focus into a very specific research question and topic

✦ no investigator can investigate all aspects of any problem on the same project, so usually research projects answer a very specific purpose in a very specific population, situation and timeframe.

FIGURE 2-1 Brainstorming Questions

Area	Questions
Values	• What are my interests and personal values? • What research topics are personally meaningful? • Have some understudied conditions that I could explore significantly affected me, my family, my friends, or my patients/clients? • Have certain health issues sparked my passion because they reflect what I consider to be an injustice?
Skills	• What knowledge and skills do I already have?
Personal growth	• What new skills do I want to develop?
Connections	• What source populations and/or data sources might be available to me through professors, supervisors, colleagues, and other personal and professional contacts?
Job and/or course requirements	• What does my supervisor or professor want me to study?
Gaps in the literature	• What information is not currently available that would make a contribution to the discipline and/or to improving health practices or policies?

2.3 Keywords:

After you have a subject, you start looking up keywords of it.

for example: you want to study child health in Africa, some of the keywords would be 'children' 'Africa' 'malaria' 'measles' 'Uganda' so basically you look up multiple words related to your main theme of the idea to give it more refinement and shape.

- Use the MeSH database (Medical Subject Heading) by the national library of medicine of the US government to identify related ideas and expand or narrow a theme.
- The MeSH dictionary is available from pubMed.org

-The MeSH database can be helpful for identifying the full extent of a research area and also for the narrowing the scope of research area.

-Once a list of keywords has been compiled, the researcher looks for the themes that emerge from them. Some topics may be easily eliminated because they do not fit the researcher's interests.

MeSH Database (by doctor): tree-like concept mapping database that helps identify related diseases and concepts altogether when doing literature search, or when establishing your own research (problem & purpose) to connect variables altogether. it also helps you later by using an international universal terminology to address your problems.

For example:

- instead of using (CVA) you'll use (stroke)

- (post) instead of (after), like post-surgery.

2.4 Exposure, Disease, Population (EDP):

- The "EDPs" form the basis for many research questions: "Is [exposure] related to [disease/outcome] in [population]?"

Example: Are exercise habits [exposure] related to the risk of bone fractures [disease] in adults with diabetes [population]?

EDP is really helpful in two major types of research approaches, which are cohorts and case-control.

Exposures could be anything, could be risk factors for diseases that are physical risk factors, environmental risk factors, biological risk factors, etc...

Example on connection between disease and exposure: contaminated water and cholera

These are examples of types of exposures →

FIGURE 2-2 Examples of Types of Exposures

Socioeconomic Status	Health-Related Behaviors	Health Status	Environmental Exposures
• Income • Wealth • Educational level • Occupation • Age • Sex/gender • Race/ethnicity • Nationality • Immigration status • Marital status	• Dietary practices • Exercise habits • Alcohol use • Tobacco use • Sexual practices • Contraceptive use • Hygiene practices • Religious practices • Use of health care services	• Nutritional status • Immune status • Genetics • Stress • Anatomy and anatomical defects • Reproductive history • Comorbidities (existing health problems)	• Drinking water • Pollution • Radiation • Noise • Altitude • Humidity • Season • Natural disasters • Population density • Travel

FIGURE 2-3 Examples of Types of Diseases

Infectious and Parasitic Diseases	Noncommunicable Diseases (NCDs)	Neuropsychiatric Disorders	Injuries
• Candidiasis • Cholera • *Escherichia coli* • Hookworm • Malaria • Syphilis • Tuberculosis	• Asthma • Breast cancer • Cataracts • Diabetes • Hypertension • Osteoporosis • Stroke	• Alzheimer's disease and other dementias • Autism • Depressive disorders • Posttraumatic stress disorder • Schizophrenia	• Bone fractures • Burns • Crush injuries • Frostbite • Gunshot wounds • Near drownings • Poisonings

Populations sometimes are referred to a group of people who share some biological characteristics or other characteristics like geographical areas.

Exposure/Disease/Population this combination of the three things usually formulates a research problem.

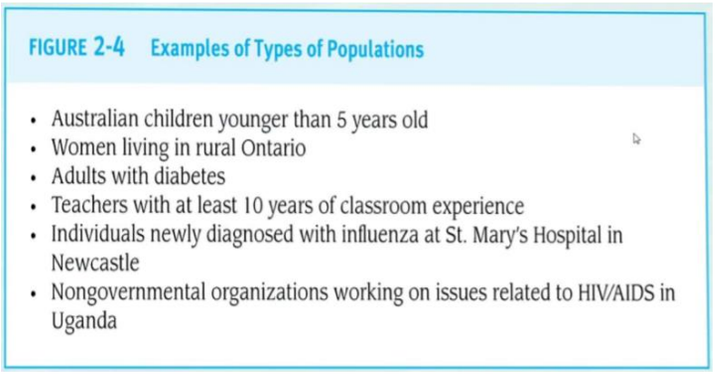


FIGURE 2-4 Examples of Types of Populations

- Australian children younger than 5 years old
- Women living in rural Ontario
- Adults with diabetes
- Teachers with at least 10 years of classroom experience
- Individuals newly diagnosed with influenza at St. Mary's Hospital in Newcastle
- Nongovernmental organizations working on issues related to HIV/AIDS in Uganda

2.5 PICOT:

- "PICOT" is often used for clinical research
  - Patient/Population
  - Intervention
  - Comparison
  - Outcome (It's sometimes called the dependent variable)
  - Timeframe

PICOT (by doctor): it's a framework used to formulate research questions and address them, usually the same framework is used as well to establish a literature search for evidence (to answer clinical questions through the literature, or to support your research arguments and to show the significance of your research topic)

- One benefit of PICOT is that it points toward the selection of key indicators that would provide evidence for the success of the intervention.

After a general research area has been identified, background reading about the topic allows the aim and scope of the research idea to be refined

So as a recap, when you do a general research area like anemia or septicemia, you do narrowing down and focusing research area and problem and looking for literature that supports your research ideas, either with or against and what has been done until this point in the area that you're exploring.

So literature search comes after this and usually it refines and focuses your research question in a better way that makes your research idea more visible and researchable with a timeframe of doing.

CHAPTER 3 : REVIEWING THE LITERATURE

3.1 Informal Sources:

- Nontechnical information from trusted sources (like the CDC and WHO) can provide helpful background on a topic.
  - Those are major reports, they are reviewed internally from the same institution, so there's no peer review. They make so many guidelines and documents and we can rely on them as researchers and we do refer to them
- Factsheets and other informal information are not part of the formal peer-reviewed scientific literature. Do not cite them in formal reports.
  - A very popular example: **Wikipedia**, it's not classified as peer-reviewed scientific literature, although it's very rich in data (it's called open resource).

-Researchers must be cautious about any claims in these files that contradict more formal sources of scientific information.

-These initial background readings can provide a foundation for understanding the more technical scientific literature that will be read later as part of through literature review.

3.2 Statistical Reports:

Statistical reports are usually published by organizations, governments or countries, and **they can be relied on**.

For example: in Jordan, we have department of statistics, they publish reports like epidemiology of distribution of certain aspects of health, or statistics about population count.

- **Examples** on statistical reports:
- World Bank world development indicators
- UN agency reports (World Health Statistics, Human Development Report, State of the World's Children)
- Annual reports from groups like the American Cancer Society and Population Reference Bureau
- Information from state and local health departments

When defining specific exposures, diseases, and/or populations of interest, it may be helpful to identify relevant statistics, such as the estimated prevalence of the exposure in a particular country, the annual global incidence of disease, or the size of a particular population.

- For regional-and country-level population measures and comparisons, the World Bank's World Development Indicators database provides information about a wide range of topics.
- Additional statistical estimates can be found in the annexes of the annual reports issued by United Nations agencies, such as the World Health Organization's **World Health Statistics**, **UNDP's Human Development Report**, and **UNICEF's State of the World's Children**.
- For information about states, provinces, counties, cities, and other smaller governmental units, contact the relevant public health department (this may be the best source of information about **vital statistics**).
- The best place to find very specific information about health-related exposures and diseases may be in published scientific articles.

3.3 Abstract Databases:

- An abstract is a paragraph-length summary of an article, chapter, or book.
- Abstract usually represents a brief description of the publication
- Use keywords to search multiple abstract databases.
  Abstract is very important in indexing beside the article title, because it's the first line search area for you as a researcher.
  Explanation: when you put only one keyword to look for all the article, you'll retrieve a large number of articles on a specific topic, but if you limit your search into title & abstract, you'll find keywords which are in that place only (so you're narrowing down your search, therefore it'll be more focused)

➔ Abstract databases allow researchers to search thousands of abstract for keywords or other terms.

➔ A careful and comprehensive search of at least one major abstract database is the most important component of a careful literature research.

In Abstract databases you should:

- Search with keywords or MeSH terms.
- Use Boolean operators: AND, OR, NOT.
- Carefully consider any limiters related to publication years or languages.

Examples of databases that are **free** to the public:

- PubMed (Which is a service of the U.S. National Library of Medicine of the National Institutes of Health, and provides access to more than 25 million abstracts)
- European PubMed Central (PMC)
- SciELO & LILACS (Central and South America)
- AJOL (Africa)

Examples of other (**usually subscription**) databases:

- **CIHAHL**: For other medical fields, but contains some medicine journals
  the criteria for indexing is different, it's technical rather than anything else, it's related to the journal and its topics, editorials and publishers, etc...
- **Embase**: A product of Elsevier company
- **MEDLINE**: previously called Medicus Index الفهرس الطبي (It's a part of PubMed, it's a very famous and old found in 1960), it's sponsored by the U.S. National library of Medicine and features only journals that have applied for inclusion and passed through a review process.
- **PsycINFO**: 3rd party database, provided by institution that are related to psychology and social sciences. So here most journals are about psycho-social health aspects.
  It's supported by the American Physiological Association (APA)
- **Web of Science**: from the company Thomson Reuters, it's an example of databases that are made by a publisher on their own.

- **EBSCO, JSTOR, Ovid, and ProQuest**: all of these are big companies that make databases, the index contains journals that are published by them, or by other publishers, and they collect all of them in one database for access (It's usually paid, not free)
- **Company-specific databases (LWW, SAGE, T&F, Wiley, others):** smaller companies which have their own databases to search for their only journals

What's the difference between Medline and PubMed?

1. Indexing in Medline is only for peer-reviewed journals, while PubMed includes documents that are not published in peer-reviewed journals, like government reports, institution reports like CDC, WHO, and many things that are not classified as journals
2. Medline uses a system of keywording that is very specific and classified, while in PubMed you can use any keyword from your mind, scientific community, etc...
3. PubMed includes books and short publications that are not classified as books nor articles

➔ **So, PubMed is more comprehensive, while MedLine is a sub-database**

3.4 Full-Text Articles:

The only way to truly understand a study is to read the full text of the article.

Where to find free full-text PDFs:

- Google scholar and other search engines
- PubMed Central and other open access repositories
- Journal websites (if the article is open access)
- Library subscription (e-journals) or interlibrary loans when a journal is not in a library's collection
- E-mail the author to politely request a copy
- Elibrary at JU, our university library has access to full texts in several databases, benefit from that.

- Google scholar → a supplement search with general search engine may be helpful for identifying additional relevant abstract, related to journals not published in English

3.5 Critical Reading:

- Read abstract
- Look at tables and figures
- Read or skim read the full article
- Review article reference list

➔ You have to take 2 values into consideration while reading:

- **Internal Validity المصداقية الداخلية :** How well was the study designed, conducted, interpreted, and reported?

- Generally speaking, internal validity relates to the structure of the article, the topic, quality of writings, time of the journal published, researchers and relevance to the topic and their experience, etc...
- More specifically, internal validity means that to what extent the quality of the paper and the way it's conducted are conducted according to scientific merits and it's being followed, and conclusions follows the introduction written by the researcher, so the conclusion is consistent with population, methodology and purpose that set up to be achieved.
- It's more like a critical judgement that is related to if the research procedures are scientifically correct and consistent with the purpose and the methods, and the conclusion reached by the researcher is appropriate to the level of data that was used.

About the internal validity, a reader should ask:
• What was the goal of the study? Were the methods appropriate for the goal? Was the main study question answered?
• Were the methods used to collect and analyze data scientifically valid? For example, did a study collecting new survey data select an appropriate sample population, recruit an adequate number of participants, use a validated questionnaire, and apply appropriate statistical tests? was the study conducted ethically? Have the authors acknowledged and discussed the limitations of the study methods?
• Do the results seem reasonable? What types of bias I the design, conduct, analysis, and interpretation of the study might have caused some of the results to be inaccurate?
• Are all of the study's conclusions supported by the study's results? If a study was attempting to answer a question about causality, does the article provide sufficient evidence to support that claim?

- **External validity (generalizability) المصداقية الخارجية :** How likely is it that the results of this study apply to other populations?
- Not all studies can be applied to all settings and all populations.
- External validity determines to what extent you can generalize your results to populations other than the population that you studied and the setting you are investigating in, but they should be similar in age group, disease, race, sex, etc...

About the external validity, a reader should ask:
• How well do the findings of this study fit with existing knowledge about the topic? Have **replication studies** in diverse populations supported the generalizability of the findings?
• For experimental studies, how likely is it that the observations from the trial would occur in everyday life outside laboratory conditions?
• To what other populations might the results apply? For example, are results from a study in Canadian men ages 30-49 likely to be applicable to Mexican men ages 30- 49, Canadian women ages 30-49, and/or Canadian men ages 50-69?

3.6 Annotated Bibliographies: • Used to track articles during-literature review

- Annotated bibliographies briefly summarize an article or report and how it relates to the proposed new project.
- it was used mostly in old days when we didn't have enough technology to help us to organize our literature findings, it's still used now adays but it's electronic.

Annotated Bibliography: we summarize findings, procedures, methods and any relevant data & information that we need from an article of interest, we summarize it as points or text either on piece of cards (Bibliographic cards) or Modern softwares like (Mendeley, EndNote or you can use reference manager in Microsoft Word)

- An annotated bibliograph includes, at minimum a full reference for the document being reviewed and a brief summary of the article or report.
- Researchers may also take notes about how a published report relates to the proposed new research project. The goal is not to replicate a document's abstract. **The goal is to summarize the content most pertinent to the new investigation.**

3.7 What Makes Research Original:

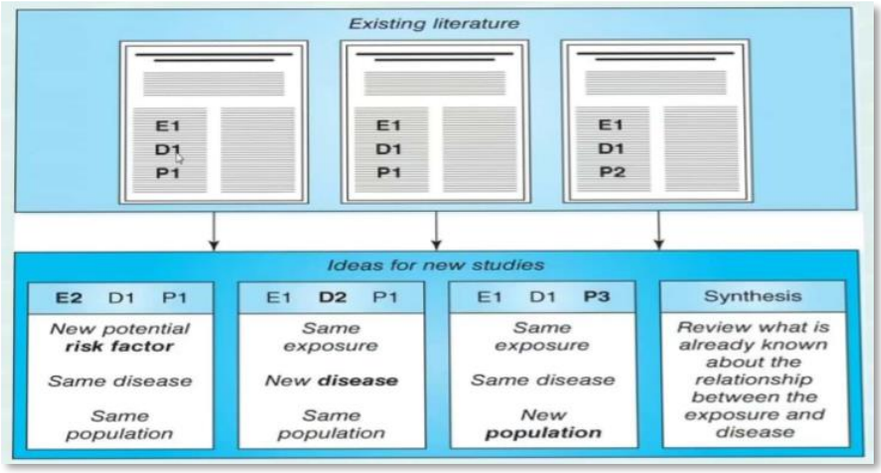
- For a research project to be considered original, it needs to have only one substantive difference from previous work: **a new exposure, a new disease/outcome, a new population, or a new perspective**.

Note:

Replication: has some changes.

Duplication: you copy the same thing, without any changes.

you can have ideas for new studies only by changing one element while the rest is the same (just see what's 'new' in the figure)



For example:

a literature review might find that several studies have shown that older adults (the population) who take 30 minutes walks several times a week (the exposure) score higher on memory tests (the disease or outcome) than adults who do not routinely walk for exercise.

A proposed new study could ask:

• Is playing table tennis (a new exposure) effective at improving memory in older adults (the same outcome and population)?

• Do older adults who walk several times a week (the same exposure and population) improve their balance (a new disease or outcome)?

• Does walking (the same exposure) improve memory (the same outcome) in children (a new population)?

CHAPTER 4: FOCUSING THE RESEARCH QUESTION

4.1 Study Approach

After identifying a general research topic, the researcher needs to develop a specific research goal and workable research plan.

To answer your questions and achieve your goals, you should first choose the type of study that you are going to do, and it can be:

- **Primary study: collect and analyze a new data,** like collecting data from a hospital records with your criteria (يعني ممكن تكون الداتا موجودة.. بس انت عليك تجمعها بالمعايير والطرق اللي بتناسبك).

- **Secondary study: analyze an existing data,** which collected for a previous study. So, you don't have a flexibility. There is no need to worry about ethical considerations of the collected data as long as you took permission from who have the copyrights

- **Tertiary study: reviews an existing literature,** two major types: Systematic reviews & Meta-analysis (it's considered complete research, not just a review. And it's expensive, so only the companies do it.

If new data will be collected, the researcher has great freedom in selecting study topics but may struggle to recruit adequate numbers of participants

FIGURE 4-2 Key Considerations

Study Approach	Key Questions to Ask
• Collection and analysis of new data	• What are possible source populations? • Will it be possible to recruit enough participants?
• Analysis of existing data	• What are possible sources of usable data files? • What questions can be explored with the available data?
• Review of the literature	• Does the researcher have access to adequate library resources? • Can the researcher reasonably expect to acquire *all* of the needed articles?

4.2 Conceptual and Theoretical Frameworks

- **A Conceptual Framework: illustrates the key relationships between EDPs** (Exposure-Disease-Population) **that will be evaluated during the study. (Much more specific in defining a relationship).**
- **A Theoretical Framework: draws on existing models in the literature to explain key relationships. (Describe a broader relationship between things (when stimulus is applied, response is expected)).**

more clarification: sometimes you may need to use some theories which may help you in your study, so you will need these frameworks. (This will be explained further in the next semester's Qualitative Research course).

- Additional note → conceptual frame work using boxes and arrows that illustrate the various relation ships that evaluated during the study

Common in nursing, social science, and educational research. But not in clinical one.

4.3 Study Goal (purpose/aim) & Specific Objectives

FIGURE 4-3 Examples of Study Goals

- To describe the incidence or prevalence of a particular exposure or disease in one well-defined population
- To assess the perceived health-related needs of a community
- To compare the levels of exposure or disease in two or more populations
- To identify possible risk factors for a particular disease in a population
- To test the effectiveness of a new preventive intervention, diagnostic test, assessment method, therapy, or treatment
- To evaluate whether an intervention shown to be successful in one population is equally successful in a second population
- To examine the impact of a program or policy
- To synthesize or integrate existing knowledge

Breaking down of the main purpose into many objectives…

First, identify ONE clear overall study goal or study question. Then, identify three or more specific objectives, aims, or hypotheses that represent steps toward answering the main study question.

4.4 Checklist for Success

FIGURE 4-4 Questions Essential to the Success of the Project

Area	Questions
Purpose and significance	• What will the study contribute? • What will be new and noteworthy about the study? • Can the importance and necessity of this project be justified? • How will the study enhance the body of knowledge in its discipline? • Who will benefit from the study besides the researcher? • How will the study help individuals and/or communities live healthier lives? • How might the study contribute to improving health practices and/or policies?
Scope and feasibility	• Is the scope of the intended project reasonable and manageable—neither too broad nor too narrow? • Can the proposed study question actually be answered? • Can the researcher answer the proposed study question?

FIGURE 4-4 (continued)

Area	Questions
Capacity and collaborators	• Does the researcher have the knowledge and skills needed to conduct the study? • Does the researcher have access to collaborators who have the expertise needed for the project? (See Chapter 5 for information on assembling a support team.)
Money and materials	• Are there adequate financial resources to conduct the study? • Does the researcher have access to equipment, space, and other physical requirements? • Given the resources available, can the researcher reasonably expect to conduct a scientifically rigorous and valid study?
Time	• Does the researcher have the time to conduct this study? • Does the researcher have the time to make this an excellent study that does not waste health resources?
Population or data	• If the plan is to collect new data from individuals, does the researcher have access to a reasonable source population and an adequate number of participants? • If the plan is to analyze existing data or to write a review paper, does the researcher have access to a reasonable existing data set and/or to an extensive library collection?
Ethics	• Will the researcher be making good use of the resources available? • Has the researcher considered the relevant ethical issues, especially those related to the collection and use of individual-level data? (See Chapter 21 for the ethical issues that should be considered.) • Is the researcher prepared to conduct culturally appropriate and scientifically rigorous research?
Target audience	• Who is likely to be interested in the findings? • [illegible] resulting paper likely to be publishable?

Good research projects are described by the acronym "FINER":

Feasible (financial and human resources)

Interesting

Novel (originality)

Ethical

Relevant

CHAPTER 5: COLLABORATION & MENTORSHIP

Professional Development

5.1 Collaborators & Consultants

Scientific research is rarely completed by one person working alone.

New investigators benefit from mentorship by several experienced researchers with different areas of expertise.

A lead researcher: the researcher who will do the majority of the work. Sometimes it is instead used to refer to the senior researcher (an experienced researcher who guides the work of a newer investigator).

It's helpful to assemble a team for a research project that is:

Scientifically valid + Ethical & Culturally appropriate + Time & Cost – effective.

Decide about co-authorship vs. acknowledgment. For example, a statistical consultants may ask to be paid by the hour to help a researcher think through analysis options as non-coauthors. These individuals who don't earn co-authorship can be thanked in the acknowledgments sections of manuscripts that benefited from their contributions.

5.2 Finding Research Mentor

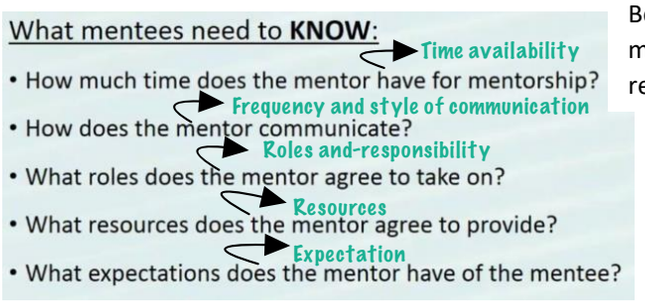
Research Mentorship: formal/informal relationship in which an experienced mentor offers professional development advice and guidance to a less experienced mentee.

A mentor is different according to the status of the new researcher (student, new researcher). New investigators seeking mentorship can identify potential advisors by:

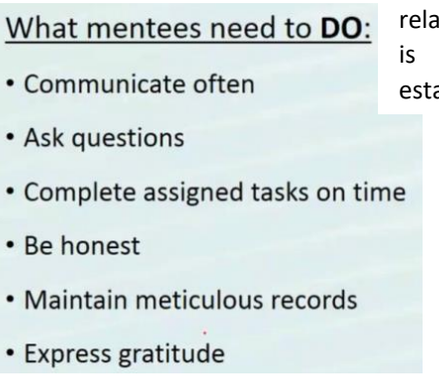
- Asking colleagues, classmates, and others about who might be a helpful mentor.
- Searching the profiles of researchers at one's own institution to see who is publishing on relevant topics.
- E-mailing potential mentors and ask to meet to discuss possible collaborations.

5.3 The Mentor-Mentee Relationship

Before entering a mentor-mentee relationship



After a mentor-mentee relationship is established



5.4 Professional Development

Don't rely on one person to provide professional development and mentoring. To establish a long-term research trajectory benefit, you should:

- **Participate in journal clubs.**
- **Become active in professional organizations.**
- **Attend and present at research conferences.**
- **Enrol in training programs.**

CHAPTER 6: COAUTHORING

6.1 Co-authorship

Most researchers start as "middle authors" before becoming a lead (first) author for the first time.

Co-authors should adhere to standards, pay attention to details, ask questions, provide variable feedback, …

Decisions about who qualifies for co-authorship should be transparent.

Decisions about coauthor ship should be made early in the research process

6.2 Authorship Criteria

ICMJE (International Committee of Medical Journal Editors) criteria for authorship in the health sciences:

All 4 criteria must be met

1- Substantial contributions to the conception and design of the study and/or data collection, analysis, and interpretation.

2- Drafting and/or critically revising the intellectual content of the manuscript.

3- Approve the final version of the manuscript to be submitted.

4- Accept responsibility for the integrity of the paper.

** **No gift authorships** (co-authorship awarded to a person who has not contributed significantly to the study) يعني مثلًا أضيف اسم صاحبي معي بالبحث وهو ما ساهم بشي منه.. أو ممكن أحتاج مصاري للبحث فبخلي ناس تساهم بالتمويل مقابل ذكر أسمائهم مع الباحثين المساهمين في البحث.. وهذا لا يجوز!

** **No ghost authors** (persons who have made a substantial contribution to the research or writing of a manuscript but are not named as authors). يعني مثلًا أحذف اسم باحث شارك معنا في بحث لدواء عشان خلافاته مع شركة معينة ما تأثر عالاعتمادية أو عالمبيعات.. وهذا لا يجوز!

6.3 Authorship order

The person who does most of the writing is often designated as the first author. The remaining authors are usually listed in order of contribution which is usually defined in terms of time dedicated to the project as well as intellectual contribution. When many co-authors with equal contributions are involved, they should be listed in alphabetical order. In prestigious journals, the senior author is often listed as the last author.

6.4 Decisions about authorship

In order to avoid last- minute debates over which individuals have made important contributions to a research project, decisions about the roles and responsibilities of each

member of the research team and who will be listed as co-author on a report, poster, or paper, as well as the order in which those person will be listed should be made as early as possible in the research processes.

قال النبي صل الله عليه وسلم:

‹ سيد الاستغفار أن يقول العبد : اللهم أنت ربي لا إله إلا أنت خلقتني ، وأنا عبدك ، وأنا على عهدك ، ووعدك ما استطعت ،

أعوذ بك من شر ما صنعت ، أبوء لك بنعمتك علي ، وأبوء بذنبي ، فاغفر لي فإنه لا يغفر الذنوب إلا أنت ›

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

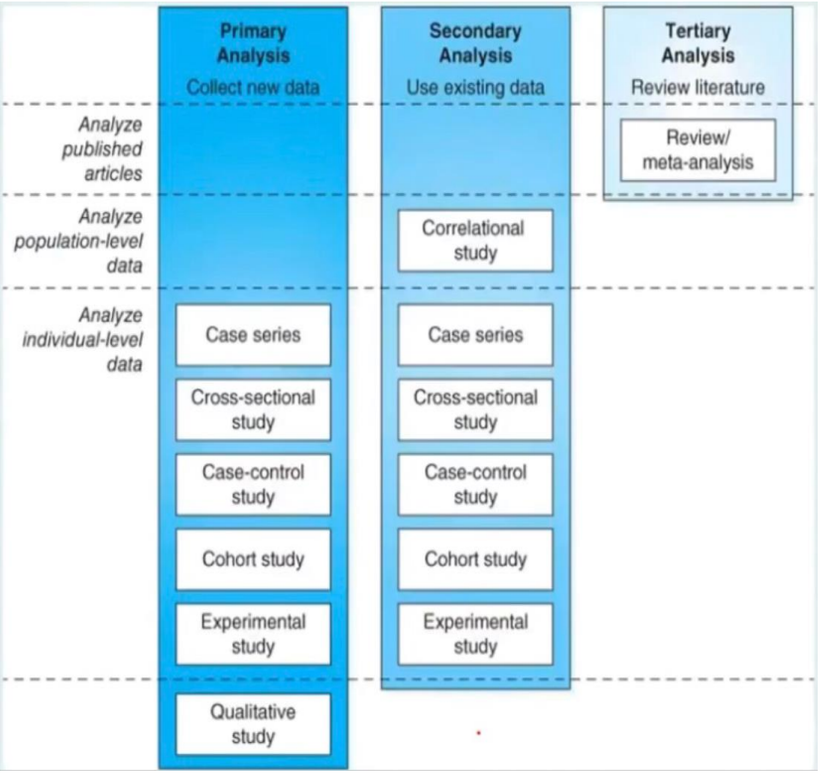
Study Approach	Goal
Case series	Describe a group of individuals with a disease
Cross-sectional survey	Describe exposure and/or disease status in a population
Case-control study	Compare exposure histories in people with disease (cases) and people without diseases (controls)
Cohort study	Compare rates 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:

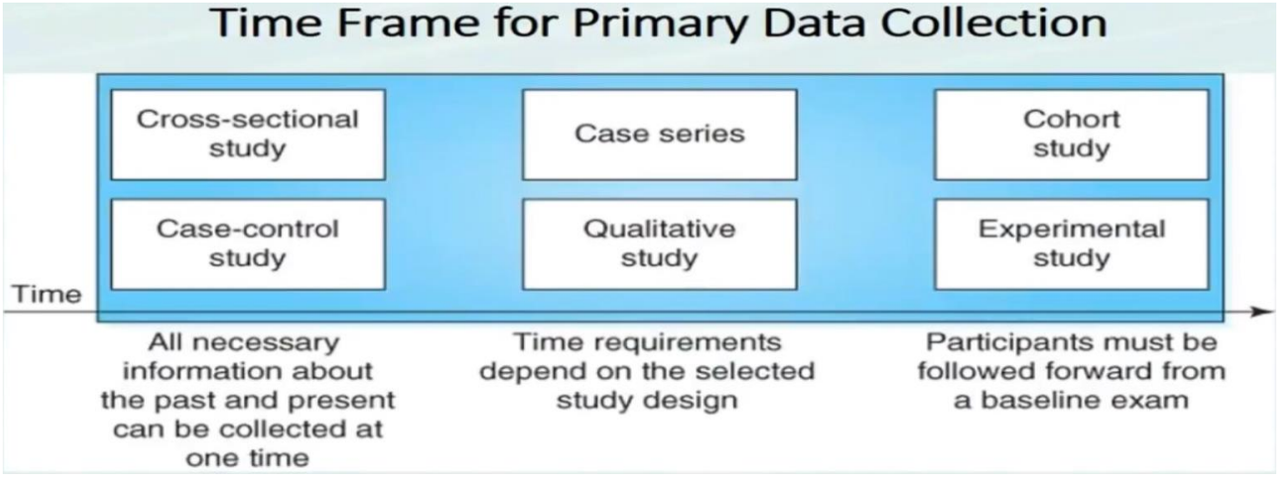
The first critical decision is whether to conduct a primary, secondary, or tertiary study. These are the study approaches, whereas what are included in the previous table are study designs (design is more specific than approach, but they can be used interchangeably). Now let's put them together >>>



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.



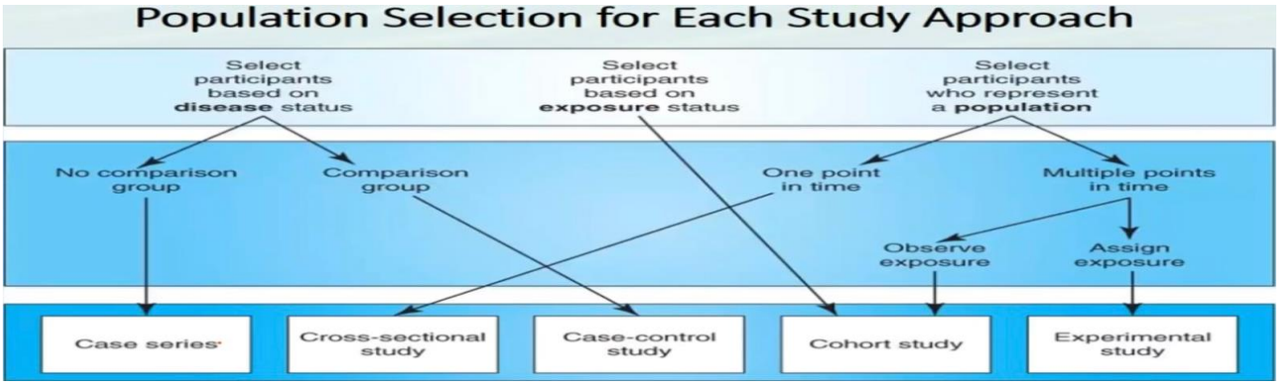
The more you go in time axis, the more effort, money, and time you need to these designs. And we will discuss each of them in details later…

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.

CHAPTER 8: CASE SERIES

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.

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

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

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

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)

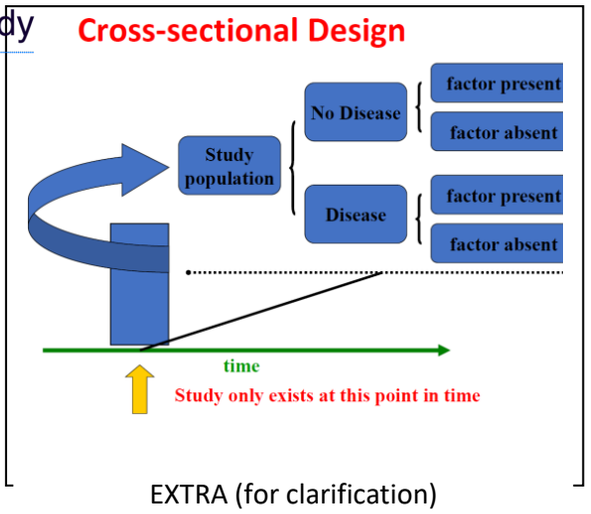
Measures the proportion of a population with a particular exposure disease.

Or

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

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

A useful video



so it ≠

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 (السببيّة/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…..

CHAPTER 10

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, and disease 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 of cases, 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 is linked to a particular case during analysis.) (1 control per 1 case)

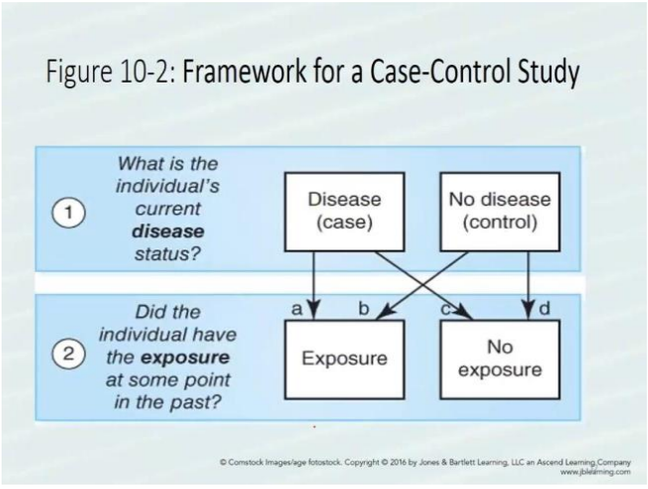
Avoid **overmatching** ,because it will be difficult to find controls who meet all the matching criteria (but if we do so we will end in a study population that is different from the general population ,so the EXTERNAL VALIDITY will be affected) → result in statistical bias that obscures the relation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

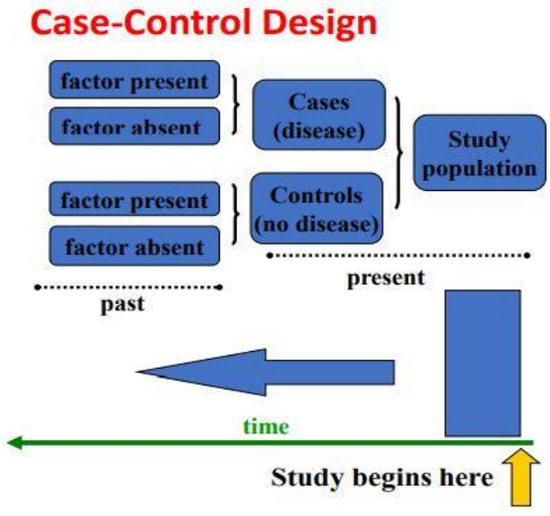
Special considerations

- Avoid ***misclassification bias*** with good case & control
- Be aware of ***recall bias,*** which occurs when cases & controls systematically have different memories of the 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 :



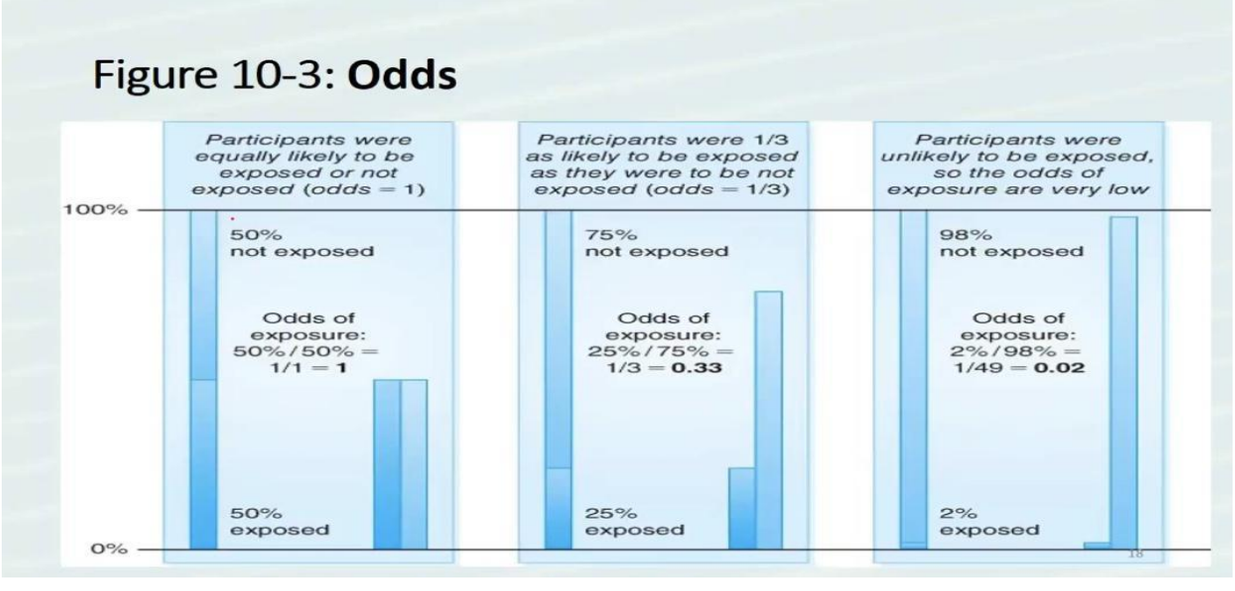


Now,

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

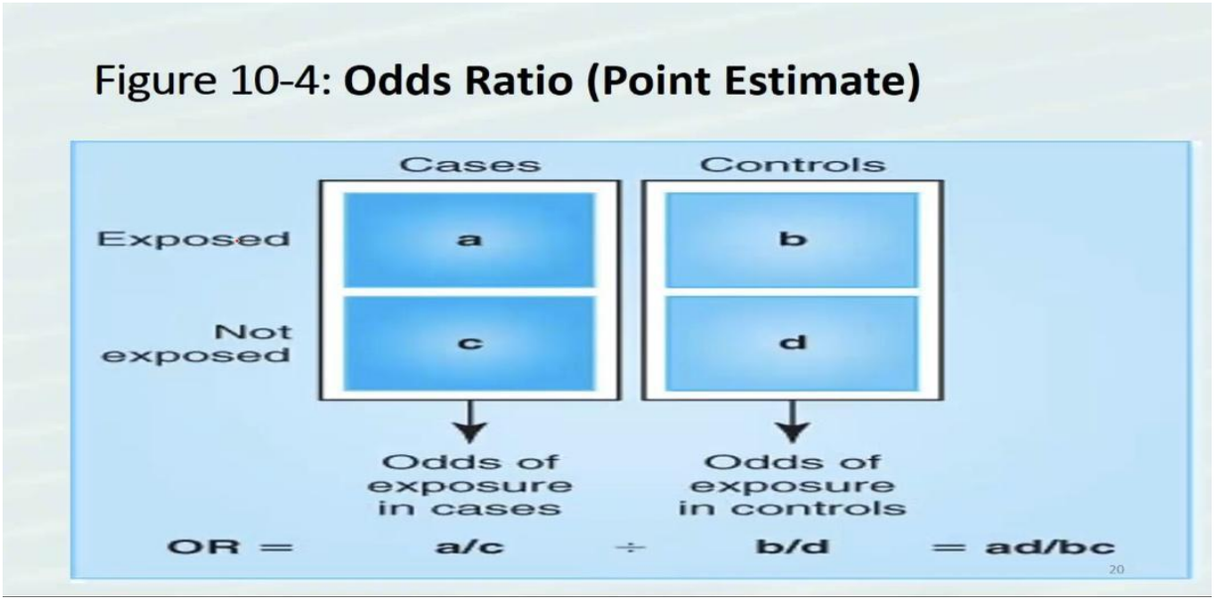
***Odds* ratio**: Compares the odds of exposure among cases to the odds among control The main measure of association in case-control study

- **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:

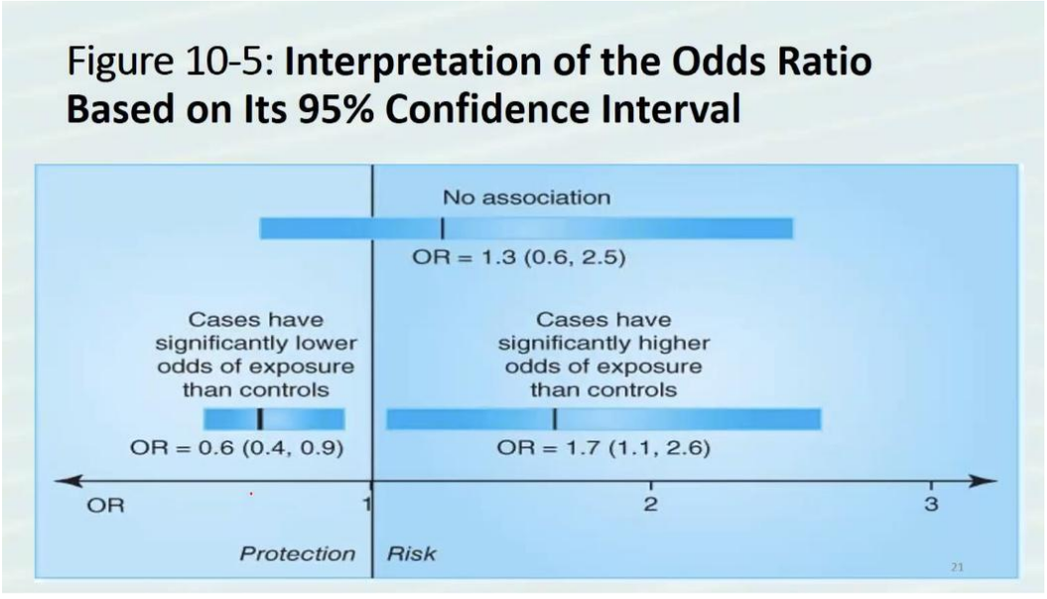


If OR:

=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.)**

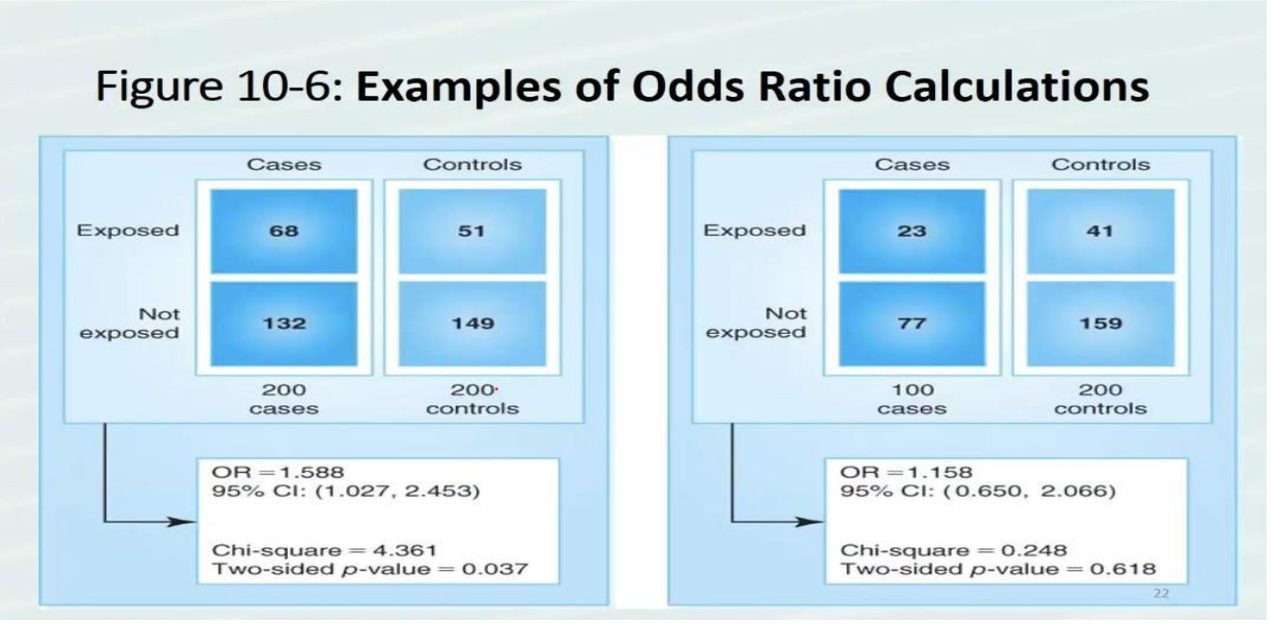


- Note : the following explanation is from the book, Dr. Jafar didn`t explain it 😊

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

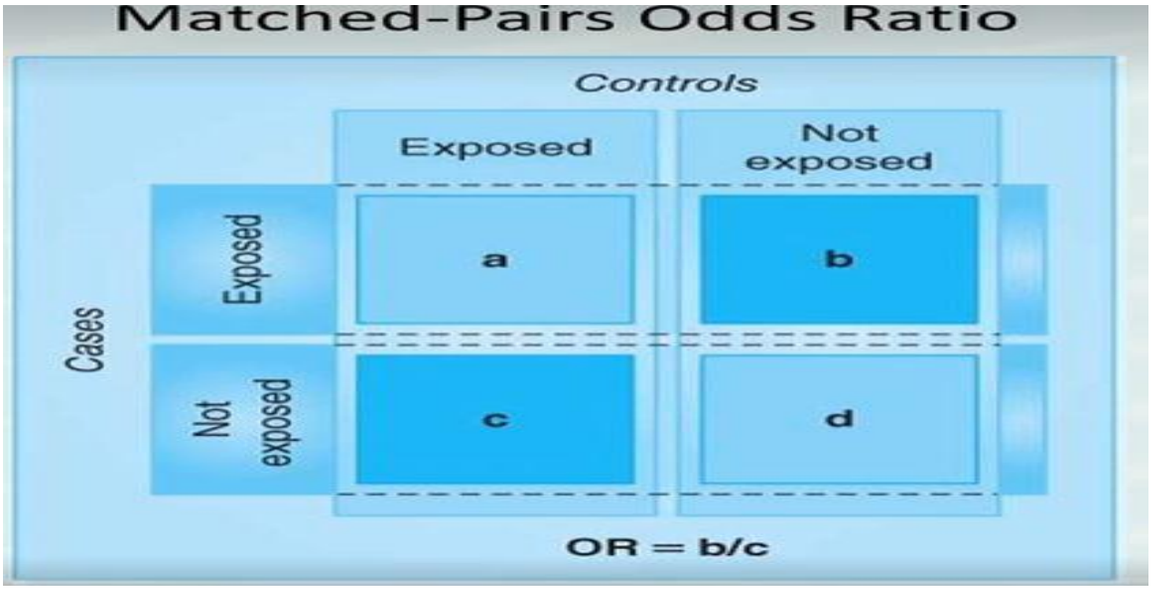
Matched case-control studies :

**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.



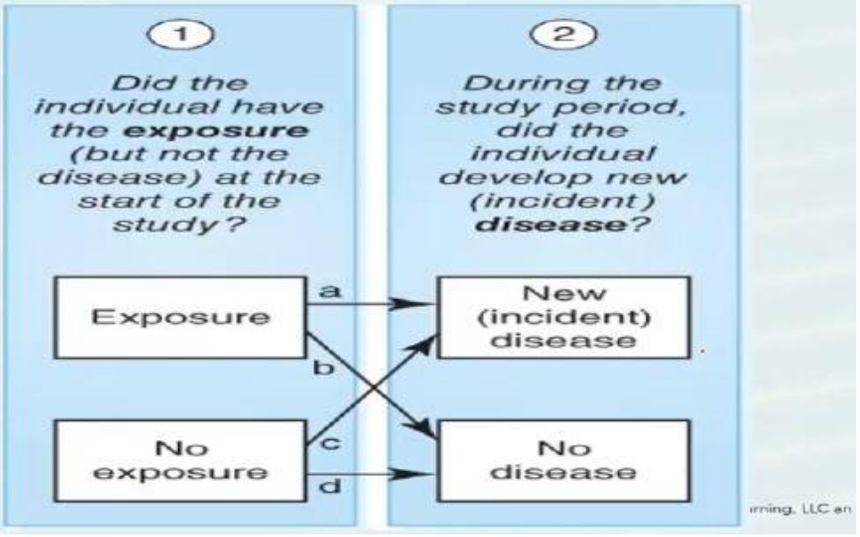
CHAPTER 11: COHORT STUDIES

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

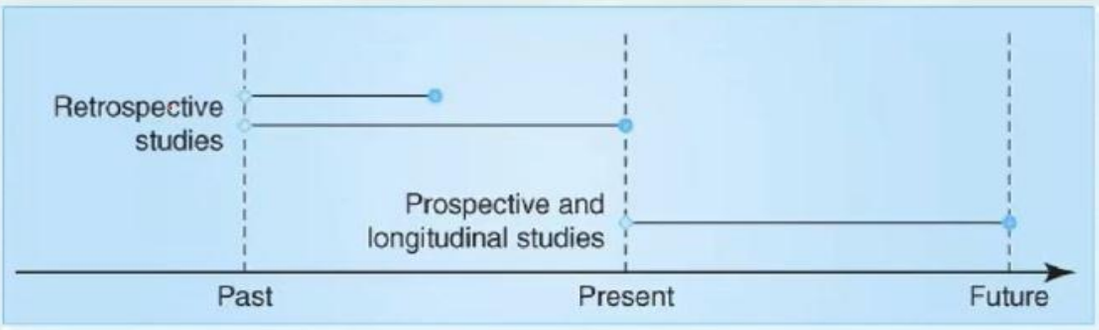
Approach	Prospective or Retrospective Cohort	Longitudinal Cohort
Objective	Compare rates of new (incident) disease over time in people with and without a particular well-defined exposure.	Follow a representative sample of a well-defined population forward in time to look for new (incident) diseases associated 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**



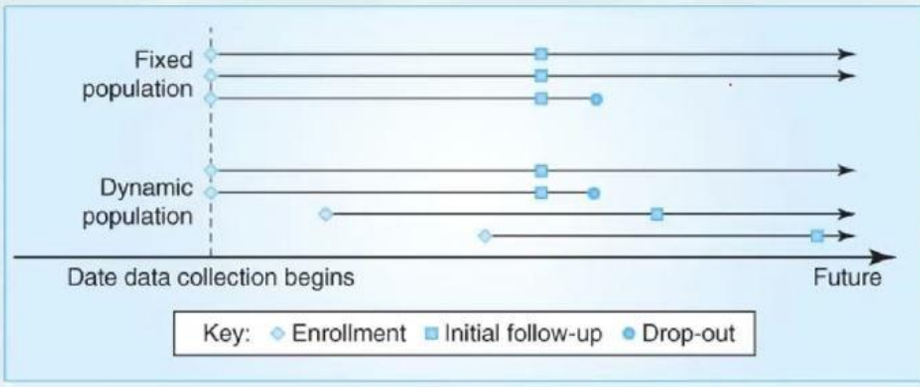
*Notice that in retrospective studies we divide the population into two groups (exposed and non-exposed -***in the past**-) and we follow up the two groups to find whether a certain disease will develop (or developed) or not

* 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 .

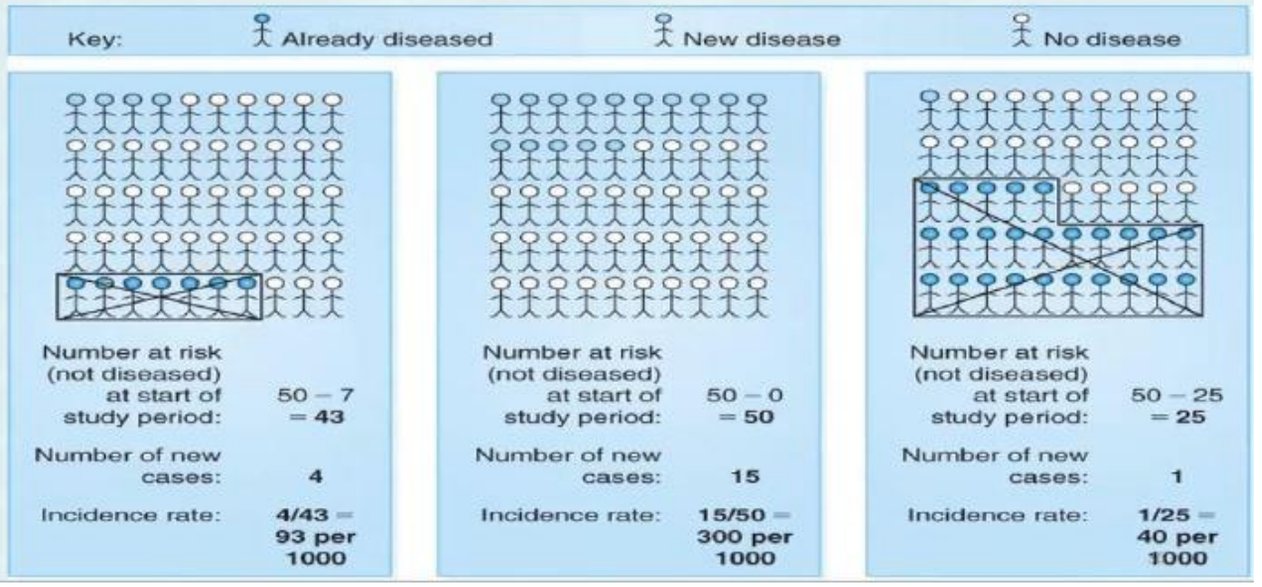
11.3 Special Considerations:

- 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

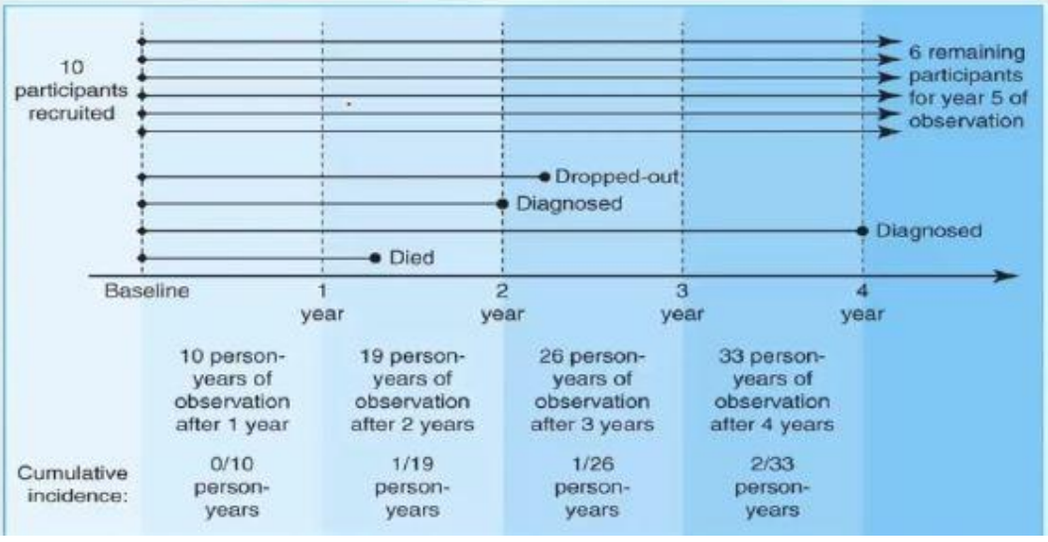
*we need to calculate I.R for exposed and unexposed groups

*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

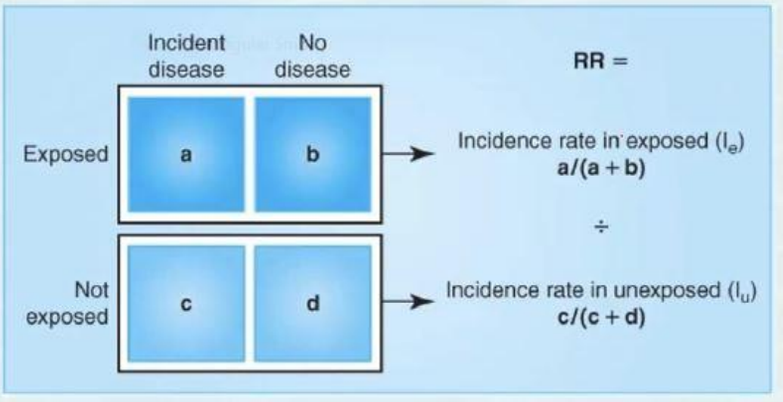
- **incidence rate ratio (RR):**

=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**

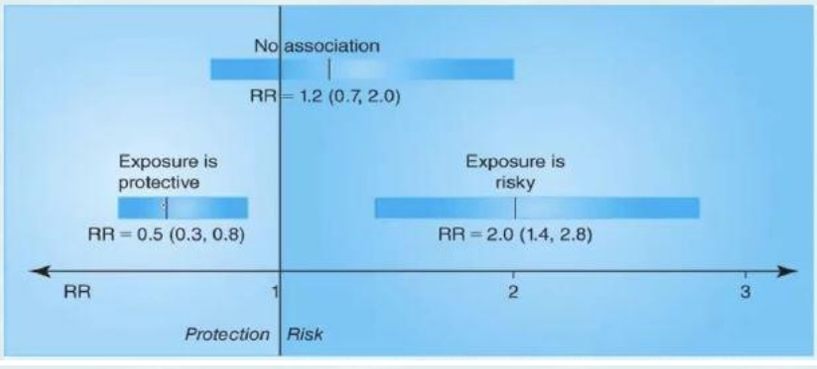
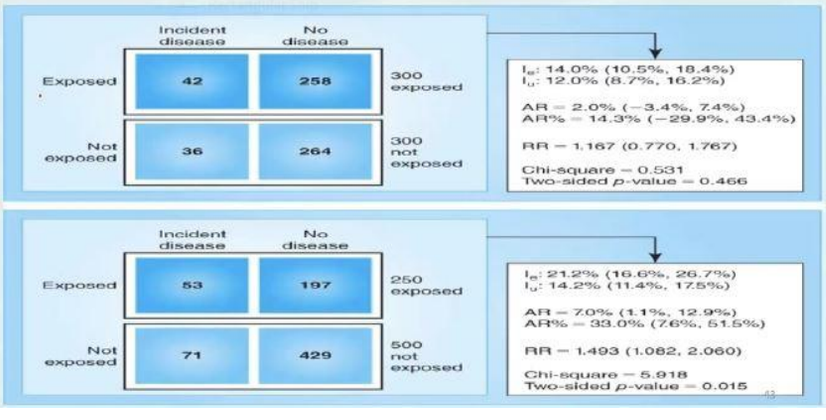


Figure 11-10: **Examples of Rate Ratio Calculations**



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

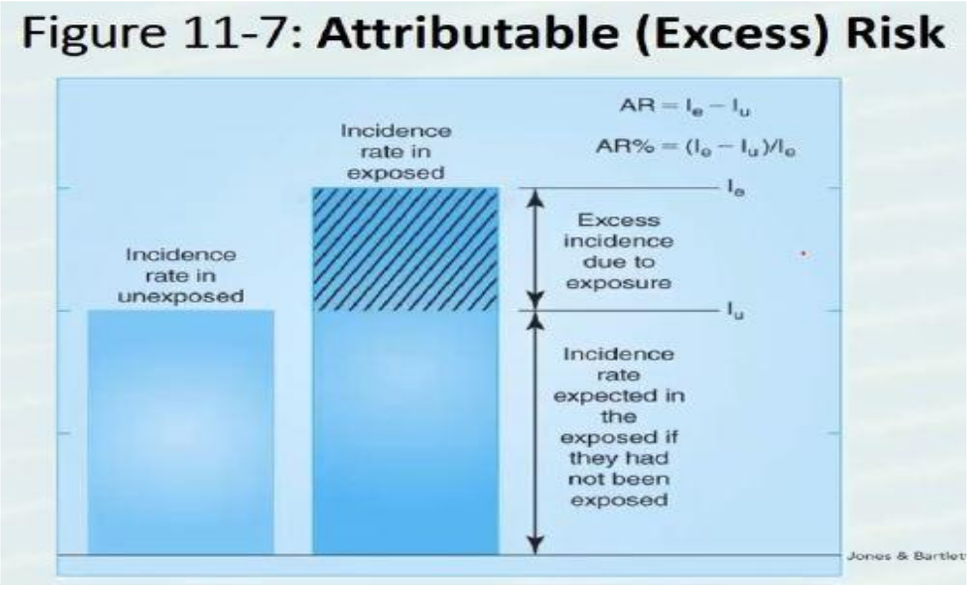
Attributable risk(AR)=Incidence in exposed-Incidence in nonexposed

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 .

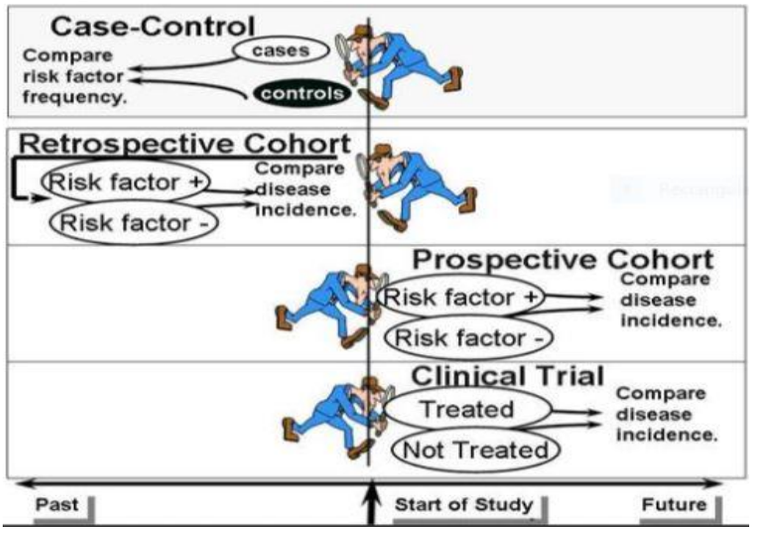
AR%= AR

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

Chapter 12: Experimental Studies

12.1 Overview:

A useful video

Experimental studies (intervention studies):

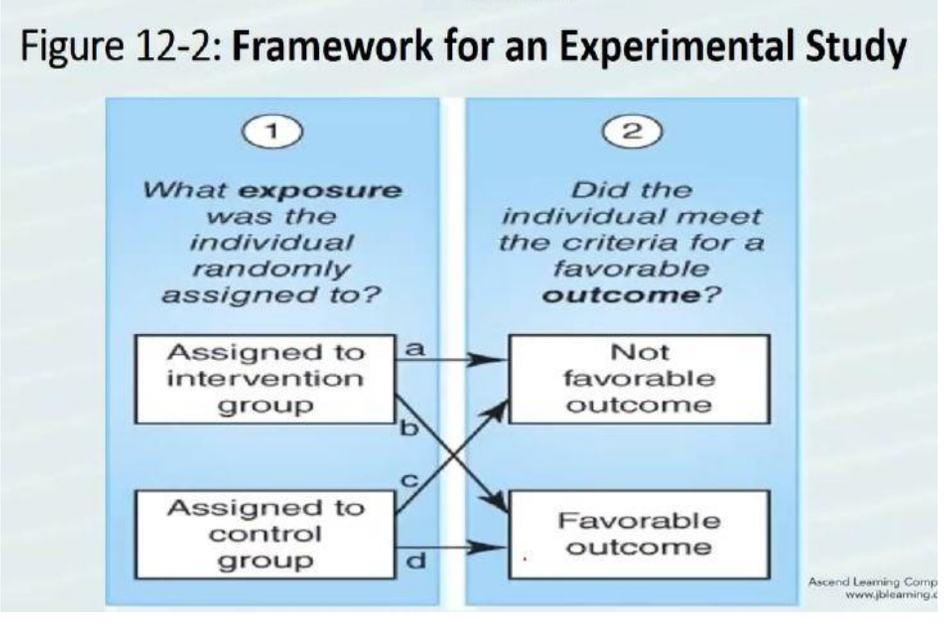
Assigns participants to intervention and control groups in order to best whether an intervention causes an intended outcome

- **Assign participants to receive a particular exposure**
- Experimental studies like Randomized controlled trials (RCTs) are the gold standard for **assessing causality**
- in **random sampling** (used in the observational studies) we select randomly the participants from the whole population ,but in **randomization**(used in experimental studies) we know exactly who are the participants we will select(corona patients ,COPD patients ,,,) but the randomization is in the selecting of control group and intervention group.
- Randomization ensures that the two study groups (control &intervention) are **comparable**(the variables are equally distributed between the two groups)

FIGURE 12-1 Key Characteristics of Experimental Studies

Objective	Compare outcomes in participants assigned to an intervention or control group
Primary study question	Does the exposure cause the outcome?
Population	Similar participants are randomly assigned to an intervention or control group.
When to use this approach	Assessing causality
Requirement	The experiment is ethically justifiable.
First steps	1. Decide on the intervention and eligibility criteria. 2. Define what will constitute a favorable outcome. 3. Decide what control is an appropriate comparison for the intervention. 4. Decide whether blinding will be used to prevent participants and/or the researchers who will assess outcomes from knowing whether a participant has been assigned to the intervention or the control group. 5. Select the method for randomizing participants to an intervention or control group.
What to watch out for	Noncompliance
Key statistical measure	Efficacy

Figure 12-2: **Framework for an Experimental Study**



12.2 Describing the intervention:

- What will the intervention be ?
- What are the eligibility criteria for participants ?
- Where and how will participants receive the intervention?
- When, how often ,and for what duration will participants receive the intervention ?

12.3 Defining Outcomes:

- **Superiority trials** aim to demonstrate that a new intervention is better than some type of control
- Researchers must carefully define what constitutes a favorable outcome for an individual participant and for the experimental study as a whole

FIGURE 12-3 Types of Success

Goal	Success
Superiority trial	The intervention is better than the control.
Noninferiority trial	The intervention is not worse than the control.
Equivalence trial	The intervention is equal to the control.

FIGURE 12-4 Examples of Favorable Outcomes

Intervention	Intended Outcome	Favorable Outcome for an Individual	Unfavorable Outcome for an Individual	Favorable Outcome for the Study Population
New diet- and exercise-based weight-loss Program	Significant weight loss	The loss of ≥10% body weight and maintenance of lower weight for ≥ 6 months	The loss of <10% body weight or failure to maintain weight loss of ≥10% or more for ≥ 6 months	The proportion of those who lose at least 10% of their body weight and maintain that loss for at least 6 months is higher in the intervention group than in the control group.
New drug therapy	Improvement of the quality of life for those with a particular disease condition	Improvement in quality of life	Failure to demonstrate improvement in quality of life	The rate of improvement in the drug therapy (intervention) group is higher than the improvement rate in the placebo (control) group, according to a carefully defined and validated set of criteria for what constitutes improvement.
New preventive vaccine	The prevention of infection	Incident infection does not occur	Incident infection occurs	The incidence of infection in the vaccinated (intervention) group is lower than the incidence of infection in the unvaccinated (control) group, as confirmed by laboratory testing.

12.4 Selecting Controls:

one commonly used type of Control is a placebo

- ***Placebo***: an inactive comparison that is similar to the therapy being test
- Some studies may compare the new therapy to some existing ***standard of care***(. we can call it metaphorically a placebo)
- ***Various*** combinations of doses and durations of an intervention can be compared using a ***factorial design***
- Participants may serve as there own controls in a ***crossover design***

FIGURE 12-5 Examples of Types of Controls

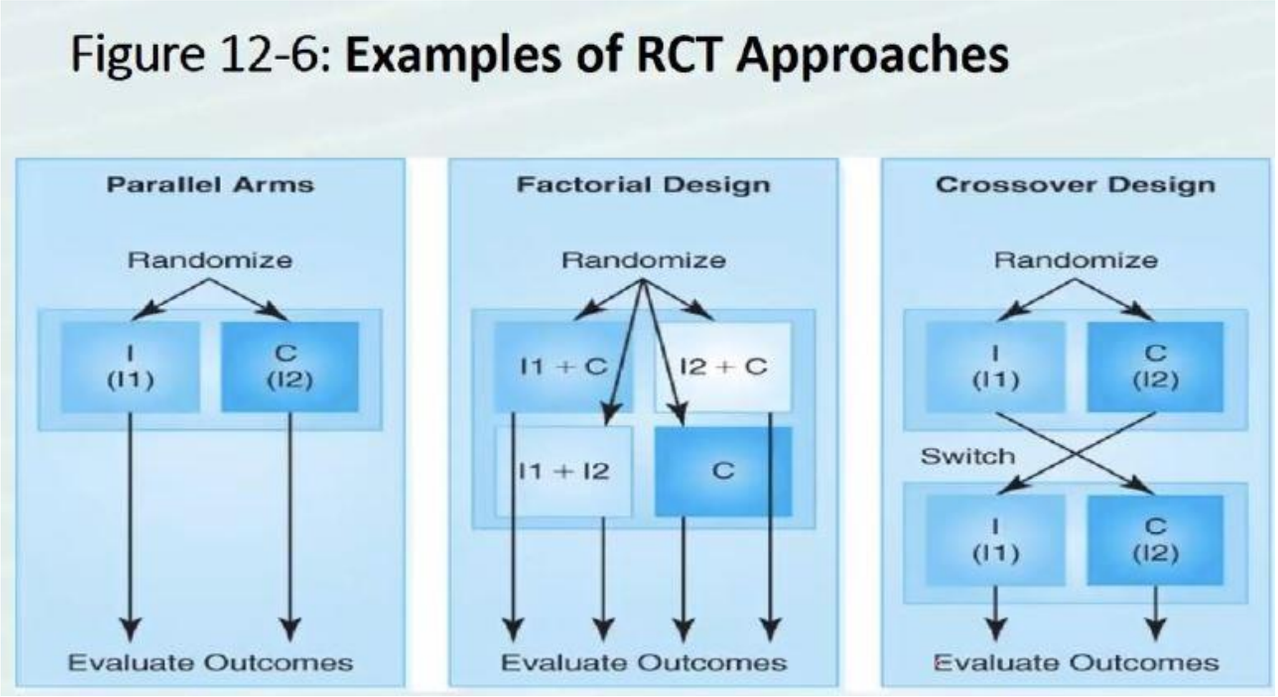
Type of Control	Active Intervention	Comparison
Placebo/inactive comparison	Active pill	Inactive pill
	Injection of an active substance	Injection of saline solution
	Acupuncture needles inserted at acupuncture points	Acupuncture needles inserted at locations in the body that are not acupuncture points (sham acupuncture)
	Some other active ingredient	An inactive substance that is indistinguishable from the active intervention in terms of appearance, odor, taste, texture, and delivery mechanism
Active comparison/ standard of care	New therapy	Current best therapy for the condition being studied
	New therapy	Current standard therapy
	New therapy	Some other existing therapy
	Current therapy plus new therapy	Current therapy alone
Dose-response	Some dose of a medication	Alternate doses of the medication
	Some duration of a therapy	Alternate durations of the therapy
No intervention	New intervention	Participants assigned to the control group are asked to maintain their usual routines.
Self	New intervention	Each participant's status before the intervention is compared to his or her own status after the intervention.
	New intervention	Each participant receives the new intervention for some duration and the comparison for some duration, preferably in a random order.

اللهم إنك عفو تحب العفو فاعف عنا

Examples of RCT approaches :

1-**Crossover Design** :every participant has the same probability to be selected in control group or intervention group

2-**Factorial Design** : we have more than one control group, or more than one intervention group, so we compare different interventions in various combinations (Ex. We give a certain drug to 3 different groups ,each of them with a different dose)



**_Hawthrone effect_ :Participants in both the active and comparison groups may change there behavior to better because they know they are being tested

12.5 blinding (masking):

**Is an experimental design element that keeps participants (and sometimes some members of the research team) from knowing whether a participant is in the active intervention group or the control group.

- we have 2 types of blinding :

1-**A single-blind experimental study**, participants do not know whether they are in an active

group or a control group.

2- **A double-blind experimental study**, neither the participants nor the researchers

assessing the participants' health status know which participants are in an active or control group.

- Blinding is intended to minimize **information bias**….. (bias) in an epidemiological study that arises due to systematic measurement error.
- There are many **types** of information bias:

1-**Reporting bias** occurs when members of one study group systematically underreport or overreport an osure or outcome.

2-**Detection bias**, also called surveillance bias, occurs when a population group that is routinely screened for adverse health conditions incorrectly appears to have a higher-than-typical rate of disease because more frequent testing enables a higher case detection rate in that population than in the general population.

3-**Observer bias** occurs when an observer (a researcher) intentionally or unintentionally evaluates participants differently based on their group membership, such as systematically evaluating cases and controls in a case–control study differently.

- Blinding prevents participants and assessors from being able to evaluate outcomes differently based on the results they expect for an exposure.
- Blinding ensures that participants in the active intervention group will not report more favorable results simply because they expect a positive outcome.
- It also keeps assessors from intentionally or unconsciously recording more favorable results for participants in the active intervention group.
- Blinding is usually possible only when all participants are assigned to similar exposures.(e.g : If participants in both the active intervention group and the control group are taking pills (of the same color, shape, size, and taste) or if both are getting injections, a blinded study may be possible. In contrast, if the active intervention is a special diet and the controls eat their usual diets, if the active group will participate in exercise classes and the controls will be on their own, or if the active intervention will include both diet and exercise components and the control only a diet plan, then a blinded study may not be possible.
- To minimize the risk of bias in studies that are not blinded, it is helpful to identify **objective** outcome measures such as laboratory tests rather than relying on **subjective** outcome measures such as participants' self-reported feelings.

12.6 Randomization:

Randomization is the assignment of participants to an exposure group in an experimental study using a chance- based method that minimizes several types of possible bias.

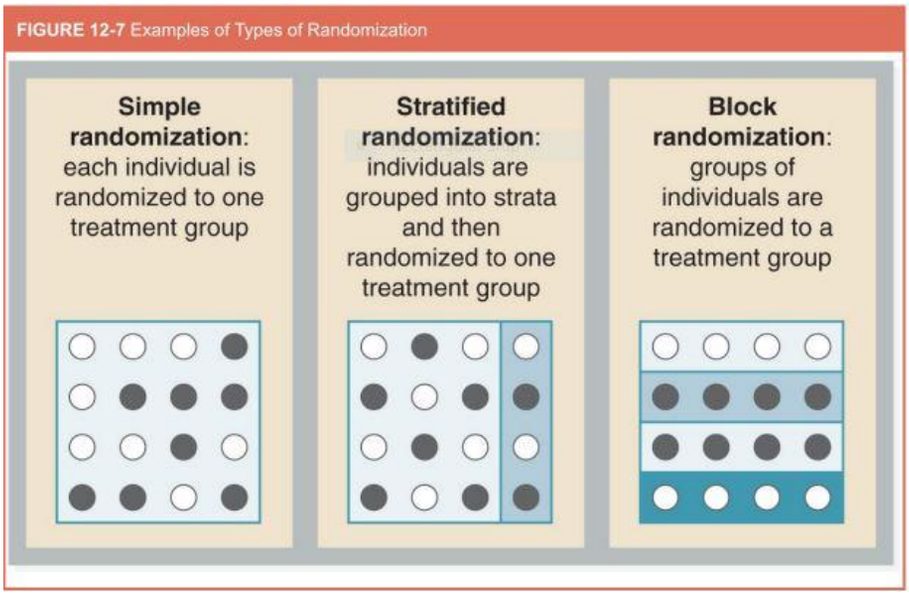
Randomization mitigates the allocation bias that might occur as a result of non-random assignment of participants to experimental study groups, such as when people with different exposure histories are not equally distributed across treatment arms.

There are many types of randomization:

1-Simple randomization is the use of a coin toss, a random number generator, or some other simple mechanism to randomly assign each individual in an experimental study to one of the exposure groups.

2-Stratified randomization is the division of a population into subgroups prior to randomly but systematically assigning each individual within each subgroup to one of the exposure groups in an experimental study.

- Stratified randomization is used when **it is important for members of certain subpopulations to be distributed evenly across the treatment arms of a trial**. For example, suppose that 75% of the volunteers for a study are female and only 25% are male. Stratified randomization can ensure that enough males are assigned to the intervention group. The list of female volunteers can be sorted into alphabetical order by last name, and then every other individual in the ordered list can be assigned to the active group. This same process can be repeated with the male volunteers. This stratified process will ensure that half of the females and half of the males are assigned to the intervention group.
  3-Block randomization is an allocation method that randomly assigns groups of people to an intervention group and other groups of people to a control group. In this method, randomization occurs at the group rather than individual level. For example, if there were 10 elementary schools in a county, schools could be randomly assigned to be intervention or control schools. All of the students in the 5 schools randomly assigned to the intervention group would receive the intervention. All of the students in the other 5 schools would be assigned to the comparison.



- Some experimental studies use nonrandomized approaches because randomization is unethical or is not feasible. A quasi-experimental design is an experimental study that assigns participants to an intervention or control group using a nonrandom method. Other than using a nonrandom method to assign participants to exposure groups, most quasi-experimental studies use methods similar to those of randomized studies. Most quasi-experimental studies use both pre- and post-intervention tests to compare the two arms of a controlled study. However, some quasi-experimental studies have no control group, and some use only a post-intervention
assessment (with or without a control group).
A natural experiment is a research study in which the independent variable is not
manipulated by the researcher but instead changes due to external forces. For example, a researcher may seek to understand the impact of a devastating tornado on the health of residents of the affected community by comparing residents in the damaged areas to residents of neighboring areas who were not directly harmed by the twister. Or suppose that a hospital announces that it will implement a new infection control policy.
group. These are not true
experimental studies because the "interventions"—a natural disaster and a policy change—are not ones that can be manipulated by a researcher, but they can be evaluated using analytic methods similar to those used for true experiments.

12.7 Ethical Considerations:

Number of issues, such as the following, must be considered before initiating an experimental study :

- The principle of equipoise states that experimental research should be conducted only when there is genuine uncertainty about which treatment will work better.
- The principle of distributive justice necessitates that the source population be an appropriate one and that the research study not exploit individuals from populations that are unlikely to have continued access to the therapy if it is found to be successful.
- The principle of respect for persons requires that all participants volunteer for a study without being unduly influenced by the prospect of being compensated for their participation. Respect also requires that all participants understand what it means to be a research subject, including the possibility of being assigned to a control group instead of the new intervention.
- The principles of beneficence and nonmaleficence require that researchers balance the likely benefits and risks of the study.
- Researchers must make careful decisions about when to use a placebo or another type of control, must put in place a system for monitoring adverse reactions, and must identify the conditions under which an experiment would be discontinued early either because the exposure proves to be risky or because the new intervention appears to be so beneficial that keeping it from the control group would be unethical

- An adverse event is a negative outcome that may be the direct result of a study-related exposure or may be a coincidental occurrence that is not directly related to the study but happens after an individual receives a study-related exposure..

FIGURE 12-8 Examples of Ethical Issues in Experimental Studies

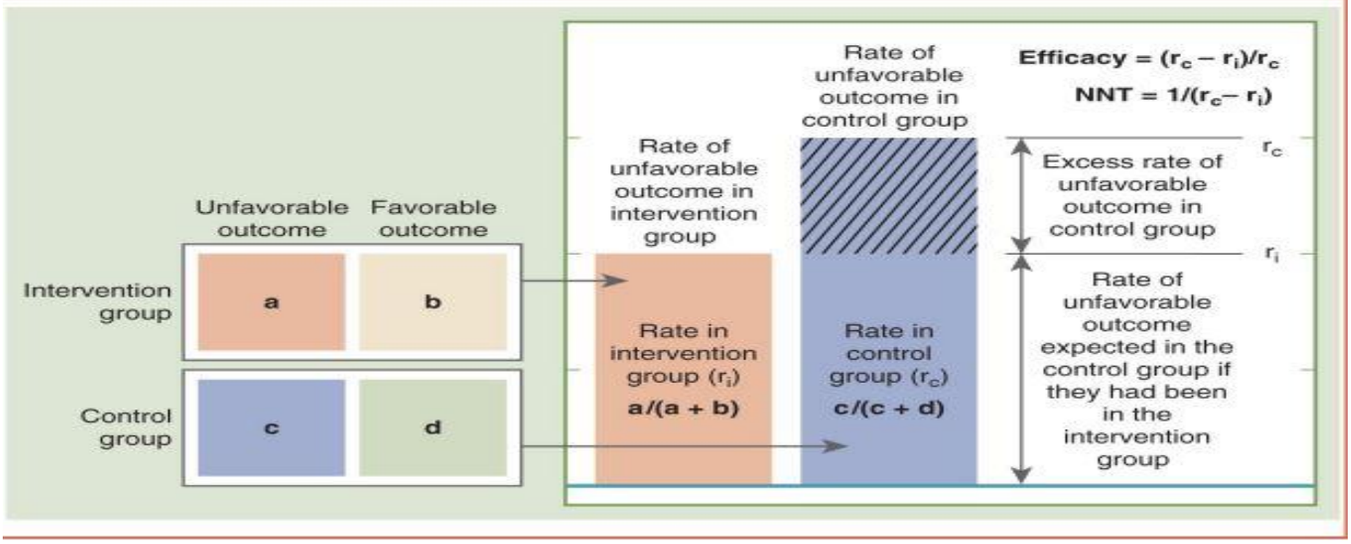
Study Stage	Examples of Questions to Ask
Study topic selection	▪ Is the study really necessary (equipoise)? ▪ Is an experimental design truly necessary?
Recruitment	▪ Is the source population an appropriate and justifiable one? ▪ Is the inducement to participate appropriate and not coercive?
Randomization	▪ Do participants truly understand that they might not receive the active intervention? ▪ Is it appropriate to use a placebo? Is it appropriate to use some other control?
Data collection	▪ How will adverse outcomes be monitored and addressed? ▪ When might an experiment need to be discontinued early?
Follow-up	▪ What happens if a participant experiences study-related harm after the conclusion of the study? ▪ Will participants have continuing access to the therapy if it is shown to be successful?

12.8 Efficacy:

Experimental studies use the statistics to quantify the impact of assigned exposure on the likelihood of having a favorable of an unfavorable outcome

Efficacy is a measure of the success of an intervention that is calculated as the proportion of individuals in the control group who experienced an unfavorable outcome but could have expected to have a favorable outcome if they had been assigned to the active group instead of the control group. A high efficacy is an indication that an intervention is successful. Efficacy typically refers to results under ideal circumstances,

FIGURE 12-9 Efficacy and Number Needed to Treat



- **The number needed to treat (NNT**) is the expected number of people who would have to receive a treatment to prevent an unfavourable outcome in one of those people (or, alternately stated, to achieve a favourable outcome in one person). A small NNT indicates a more effective intervention. If a drug intended to prevent stroke has an NNT of 5, then 5 people have to take the drug for 1 year (or some other specified time period) to prevent 1 of the 5 from having a stroke. If the drug has an NNT of 102, then 102 people have to take the drug to prevent 1 of the 102 from having a stroke.
- **The number needed to harm (NNH**) is the number of people who would need to receive a particular treatment in order to expect that one of those people would have a particular adverse outcome. A large NNH indicates a safer intervention. NNT and NNH are often used for cost-effectiveness analysis.

- **Effectiveness** is calculated with the same equation as efficacy but refers to results obtained under real-world, less-than-ideal conditions. For example, in a real-world setting, some participants might skip some doses of an

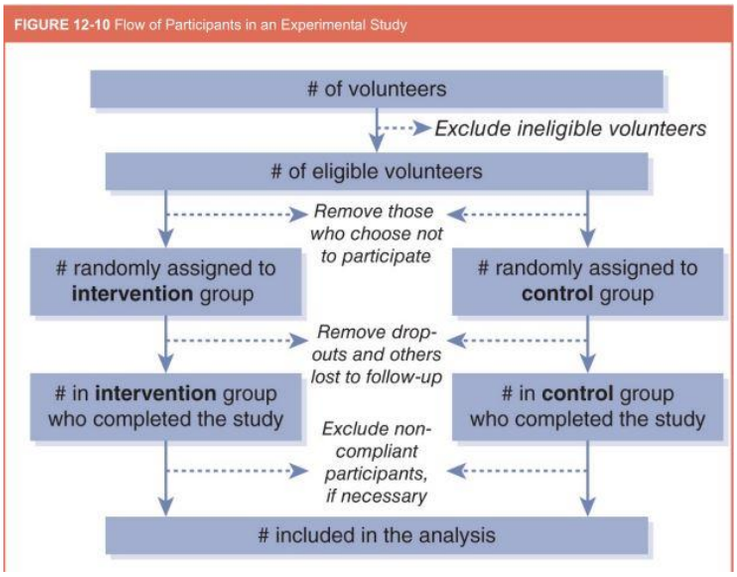
experimental drug, or they might not take the doses at the exact specified times, or they might not store the pills at the ideal temperature.

- **Efficiency** is an evaluation of the cost-effectiveness of an intervention that is based on both its effectiveness and resource considerations.

- Analysis for experimental studies typically uses either a **treatment-received approach** or **a treatment-assigned approach**.

  A treatment-received analysis of experimental data includes only the participants who were fully compliant with their assigned intervention or comparison protocol. Treatment-received analysis allows for the calculation of efficacy

  A treatment-assigned analysis (or intention-to-treat analysis) includes all participants, even if they were not fully compliant with their assigned protocol. Treatment-assigned analysis is better at measuring real-world (rather than ideal-world) effectiveness.



12.9 Screening and Diagnostic Test:

The goal of some studies of screening or diagnostic tests is to compare two assessment that are supposed to measure the same thing

A useful video

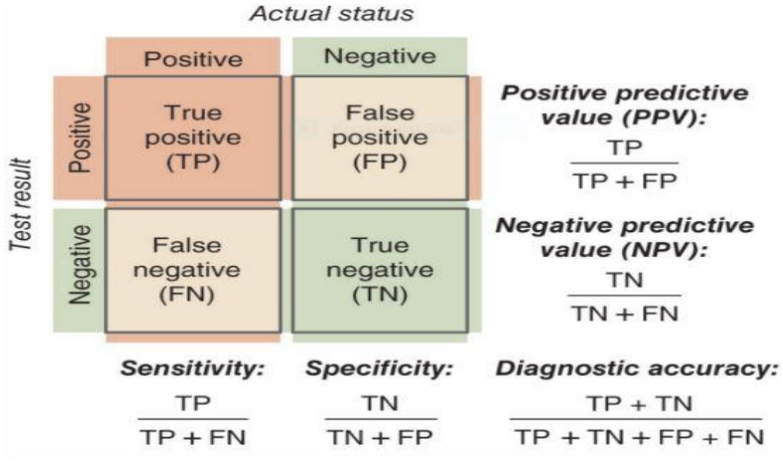
- A good test will have a value near 100% for the following **four** calculations :

1-**The sensitivity**, or true positive rate, is the proportion of people who actually have a disease (according to the reference standard) who test positive using the new test.

2-**The specificity**, or true negative rate, is the proportion of people who do not have the disease who test negative with the new test.

3-**The positive predictive value (PPV)** is the proportion of people who test positive with the new test who actually have the disease (according to the reference standard).

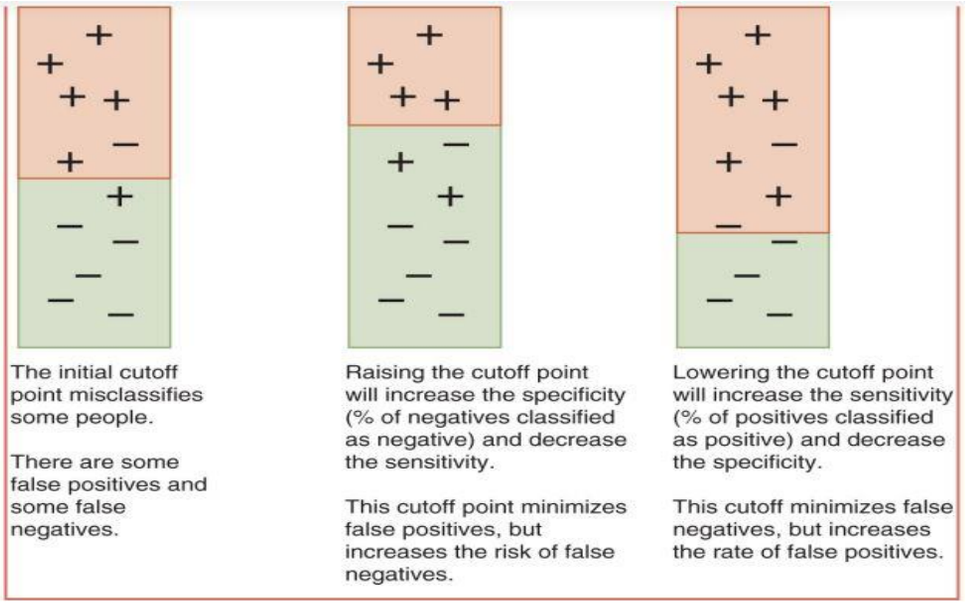
4-**The negative predictive value (NPV)** is the proportion of people who test negative who actually do not have the disease



For tests with a flexible cutoff point for defining positive and negative test results, there is always a trade-off between sensitivity and specificity (Figure 12-12). Increasing the sensitivity decreases the specificity. Increasing the specificity decreases the sensitivity. Consider the use of systolic blood pressure as a sign of hypertension

.Suppose that the cutoff for being classified as having clinically high blood pressure is reduced from 160 mm Hg to 140 mm Hg. The sensitivity of the test will increase, because a higher percentage of people with hypertension will be classified as hypertensive. The specificity will decrease, because a lower percentage of people without hypertension will be correctly classified as not being hypertensive

Three other measures are also commonly used for screening tests:
1- diagnostic accuracy → is the percentage of the participants who where either turn positive or true negative
2- positive likelihood ratio → examine whether anew test is good at predicting the presence of disease
3- negative likelihood ratio → examine whether a new test is good at predicting the absence of disease



*Chapter 13 is NOT included in the course

Chapter 14: Correlational Studies

14.1 Overview:

Useful videos : 1

2

- A ***correlational study/ecological study/aggregate study*** uses population-level data to look for associations between two or more group characteristics.
- No individual-level data are used.

Ex: we have 1000 participants , 200 of them have a stroke ,and the rest are not(we do not have control/case groups ,the same group is already divided) .So we look at the characteristics of each of the two groups.

FIGURE 14-1 Key Characteristics of Correlational (Ecological) Studies

Objective	Compare average levels of exposure and disease in several populations
Primary study question	Do populations with a higher rate of exposure have a higher rate of disease?
Population	Existing population-level data are used; there are no individual participants.
When to use this approach	The aim is to explore possible associations between an exposure and a disease using population-level data.
Requirement	The topic has not been previously explored using individual-level data.
First steps	1. Select the sources of data that will be used. 2. Decide on the variables to include in the analysis.
What to watch out for	The ecological fallacy Limited publication venues
Key statistical measure	Correlation

14.2 Aggregate Data:

- At least two population-level indicators must be available for each population (defined by place or time).
- These "exposures" and "outcomes" must be measured similarly in all populations being compared.

FIGURE 14-2 Sample Data Table

Population	Exposure 1	Outcome 1
A	48.2	14.1
B	65.1	17.0
C	37.8	14.9

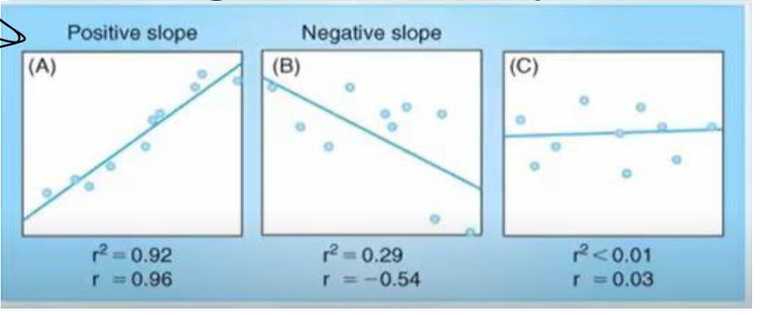
14.3 Correlation:

*the type of equation we use to represent the correlation depends on the level of the measurements of the variables (continuous, categorical ….)

* r^2 represent the significance

a- A positive slope → shows that higher levels of exposure are associated with higher rate of disease
b- A negative slope → snows that higher level of exposure are associated with lower rate of disease

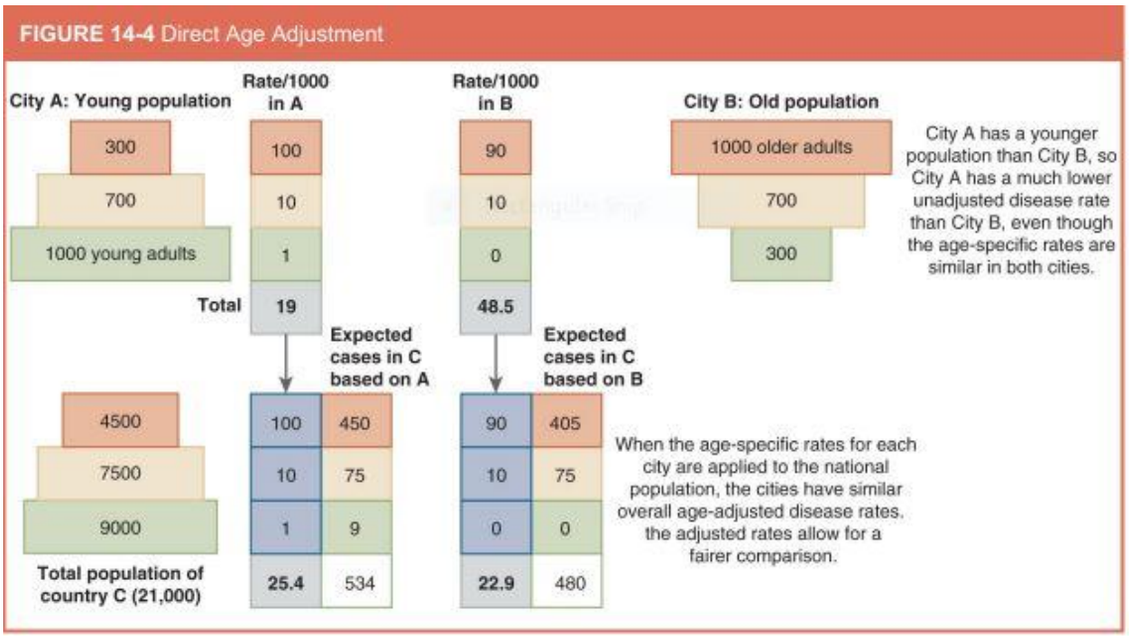
- For a two-variable analysis, plot each population on a scatterplot with the "exposure" on the *x*-axis and the "outcome" on the *y*-axis.
- A best-fit line defines the correlation (*r*) between the two variables.
- Use linear regression to fit more complex models of



14.4 Age adjustment:

- Use age-adjustment to more fairly compare two populations with very different age distributions.
- ***Direct age adjustment*** requires knowing age-specific rates of exposure and/or disease as well as the age distributions of the populations being compared.
- ***Indirect age adjustment*** does not require age-specific rates.

- The ***ecological fallacy*** is the incorrect attribution of population-level associations to individuals.



كان ابن مسعود ‹رضي الله عنه› يقول:

‹إني لأكره أن أرى الرجل فارغا، لا في عمل آخرة ولا في عمل دنيا›

Chapter ~~22~~ 15: Systemic Reviews and Meta-Analysis

A **systematic review** is the careful compilation and summary of all publications relevant to a particular research topic, and a **meta-analysis** creates a summary statistic for the results of systematically identified articles.

22.1 Overview:

- Most scientific research projects seek to identify new findings derived from a single study population, but the goal of tertiary analyses is to engage in the scholarship of integration.
- **Synthesis research:** integrates existing knowledge from previous research projects. The common types of synthesis research in the health sciences include narrative reviews, systematic reviews, and meta-analyses.
- **The goal of a review article** is to synthesize what is already known about a topic by connecting previous studies and offering new interpretations of their contributions to scientific knowledge.
- However, review articles have limitations. Not all journals publish review articles as they are regarded as exhibiting less originality than other types of research.
- **All reviews require:**
  - An extensive search of the literature.
  - The extraction of key information from relevant articles.
  - The clear and concise presentation of this information.

➔ They all have the same objective and population.

FIGURE 22-1 Key Characteristics of Reviews and Meta-Analyses

Approach	Narrative Review	Systematic Review	Meta-Analysis
Objective	Synthesize existing knowledge	Synthesize existing knowledge	Synthesize existing knowledge
Primary study question	What conclusions about this topic are supported by previous studies?	When all previously published studies on this topic are examined, what conclusions can be drawn?	When the results of all previously published studies on this topic are merged, what is the summary statistic?
Population	Published literature	Published literature	Published literature
When to use the approach	The goal is to describe a new perspective on a topic that can be supported by the existing literature.	The goal is to compare the findings of previous studies on a well-defined topic.	The goal is to summarize previous findings using pooled statistics.
Requirements	The researcher has excellent library access. The researcher has a unique perspective on the topic.	The researcher has excellent library access. The researcher can obtain every relevant article.	The researcher has excellent library access. The researcher has strong quantitative skills.
First steps	1. Decide what story the article will tell.	1. Decide on the specific objectives of the review. 2. Select the search methods that will be used to find potentially relevant articles. 3. Select inclusion and exclusion criteria for articles.	1. Decide on the specific objectives of the review. 2. Select the search methods that will be used to find potentially relevant articles. 3. Select the inclusion and exclusion criteria for the articles. 4. Decide how to assess the quality of the studies. 5. Decide how the results of the studies will be combined into one summary statistic.
What to watch out for	Limited publication venues	Publication bias	Studies that cannot be fairly compared
Key statistical measure	No statistics are required.	No statistics are required, but providing some results from included studies may be helpful.	Summary measures for included studies must be reported.

22.2 Selecting a topic:

- ❖ Each review needs to have an appropriate scope. So, when starting a tertiary analysis, the most important decision is the selection of a topic that is narrow enough that all the relevant publications can be acquired.
- ❖ Most successful reviews have more than just a few articles & less than hundreds of articles.
- ❖ The topic may need to be modified if a preliminary literature search does not yield an appropriate number of articles. If an initial search of an abstract database yields only 8 possibly relevant articles, the topic probably needs to be expanded. If a search produces 352 articles, the topic needs to be narrowed to a more specific disease condition, a smaller geographic area, or a reduced scope.

➔ **For example**: a review of risk factors for cardiovascular disease would be cumbersome. A very long book would be required to cover all the identified risk factors, and an article-length summary would provide such a superficial level of information that it would not be useful. There is a greater likelihood of success for a review article with a narrower scope—one that limits the types of risk factors, the particular cardiovascular diseases, and the population groups included in the analysis.

22.3 Library access:

- ❖ No review article can be written without excellent library access because **every relevant article** must be identified and obtained during a systematic review
- ❖ The article acquisition process usually requires access to a university library that allows affiliates to make numerous interlibrary loan requests.
- ❖ Check with an institutional librarian about policies and prices for accessing articles that are not part of the library's collection.
- ❖ The researcher must also prepare to maintain a meticulous system for tracking articles that have already been acquired, those that have been requested but not yet received, and those that need to be requested.

22.4 Narrative Reviews:

- ❖ Narrative reviews tell a story about a topic using evidence from the literature to support the "plot".
- ❖ A narrative review might summarize important clinical aspects of a disease or summarize the epidemiological profile of a well-defined population.
- ❖ Narrative reviews must be carefully organized by theme, methodology, chronology, or some other guiding principle, because they are intended to convey a perspective and not merely compile facts.
- ❖ A narrative may also be appropriate when the researcher has developed a unique conceptual framework or theory that can be illustrated with examples from the literature.
- ❖ A narrative review works best when the researcher has a unique perspective on a topic and/or a particular expertise in the field.

➔ However, narrative reviews are becoming less common as editors and reviewers push for the use of systematic methods. So, Researchers must be prepared to justify their selection of a narrative approach.

22.5 Systematic Reviews:

Systematic reviews use a predetermined & comprehensive searching & screening method to identify relevant articles while minimizing bias that occur when researchers handpick the articles they want to highlight.

➔ Therefore, after the identification of a focused study question, the most important decisions in a systematic review **are the selection of keywords and inclusion criteria.**

➔ The goal is to craft a search strategy that identifies all the articles ever published on the narrow, well-defined area covered by the review.

16

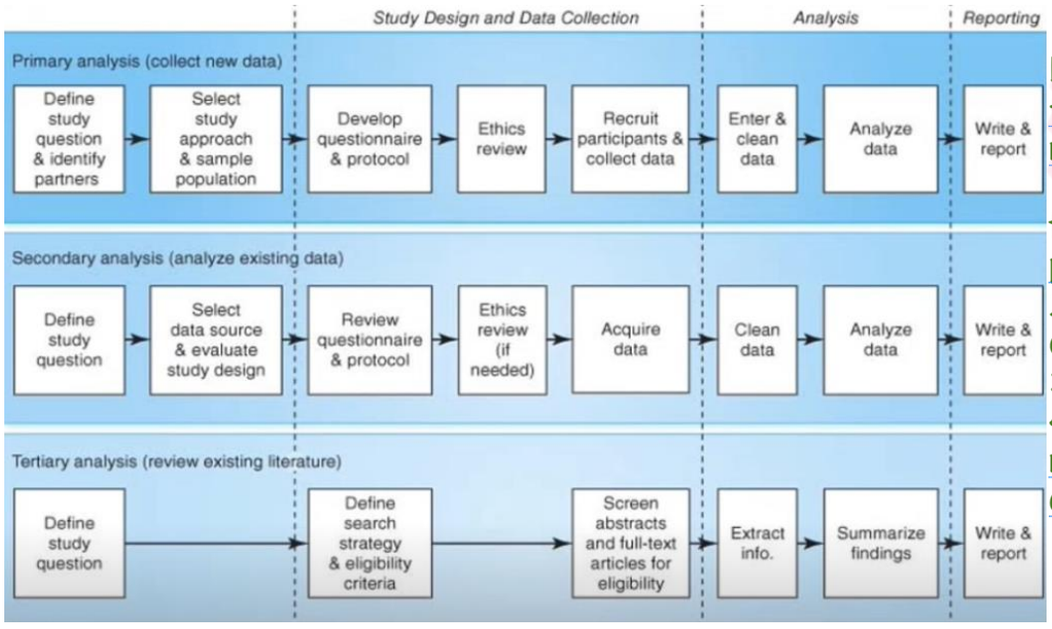
Chapter 15 : Research Protocols

hand book describing all the actions that will be taken during the implementation of the research plan

15.1 Overview of Research Plans by Study Approach

A research protocol is a detailed written description of all the processes and procedures that will be used for data collection and analysis.

All types of studies -primary, secondary, and tertiary- require a research protocol to be created prior to the collection & analysis of data. (as shown in the following figure).



For the collection of new data from individuals, the researcher needs to:
1. Identify an appropriate way to sample and recruit participant
2. Develop a questionnaire and other data tools
3. Select methods
4. Prepare an application for a research ethics review committee

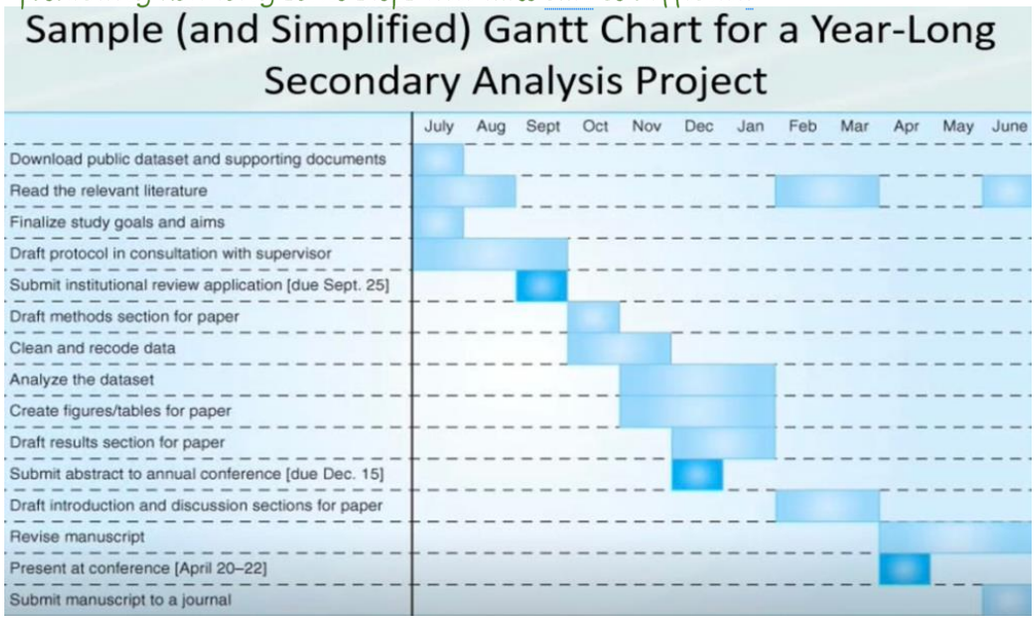
15.2 Research Timelines

It is helpful to create a project calendar (such as a Gannt Chart) that specific critical deadlines & other steps toward successful and timely completion .

Some flexibility is required **(because predicting how long some steps will take can be difficult).**

GANTT CHART → very helpful for visually displaying the research time line.

Note: the internal due dates set by the research team will need to be some what flexible because predicting how long some steps will take can be difficult.



15.3 Research Responsibilities:

The roles & responsibilities accepted by each collaborator should be defined early in the project.

Institutions usually require one person to be designated as the primary investigator (PI) with special responsibilities for ensuring that:

a) the protocol is followed.

b) the budget is maintained.

c) adverse outcomes are promptly reported.

Primary investigator → is the person doing the greatest amount of work on the project .
Many institutions allow only senior employees to serve as official institutional PIs. For example, some universities require a professor to be listed as the PI on any research project that involves human subjects, even if a student is taking the lead role in the conduct of the project

15.4 Writing a Research Protocol

A research protocol should describe the exact procedures that will be used for every step of the research process.

A strong protocol provides enough detail that the study could be replicated by other research teams.

Ideally, a protocol should:

1.Fully describe all the procedures that will be used for data collection and analysis.

2.List the anticipated dates of completion for each of the steps in the research process.

3.Provide details about the responsibilities of each member of the research team.

4.Describe the mechanism for updating any part of the research plan if revisions arise after approval of the initial protocol.

15.5 Preparing for Data Collection

Data collected should only be initiated **after** a data management plan is in place.

History of clinical research and its ethics:

Here are some recent events in medical research that formed and shaped our current guidelines, regulations and system:

1932–1972: Tuskegee Syphilis Study

The most notorious example in the United States of prolonged and knowing violations of the rights of a vulnerable group of research participants.

That study, conducted under the auspices of the U.S. Public Health Service (PHS) at Tuskegee Institute (now Tuskegee University) in Tuskegee, Alabama, was originally projected to last six months but spanned 40 years—from 1932 to 1972. The purpose of the study was to determine the effect of untreated syphilis in black men. The men in the study were never told that they had syphilis.

This research used disadvantaged, rural black men to study the course of an untreated disease.

The men were offered free examinations and medical care but were not informed of their disease, that they were participating in research, or that the research would not benefit them.

Further, in order not to interrupt the project, participants were deprived of demonstrably effective treatment long after such treatment was discovered and had become generally available.

1939–1945: Nazi Experiments During World War II

Although not the first example of harmful research on unwilling human participants, the experiments conducted by Nazi physicians during World War II were unprecedented in their scope and the degree of harm and suffering to which human beings were subjected.

"Medical experiments" were performed on thousands of concentration camp prisoners and included deadly studies and tortures such as injecting people with gasoline and live viruses, immersing people in ice water, and forcing people to ingest poisons.

In December 1946, 23 physicians and administrators, many of them leading members of the German medical hierarchy, were indicted before the War Crimes Tribunal at Nuremberg for their willing participation in the systematic torture, mutilation, and killing of prisoners in experiments.

Despite the arguments of the German physicians that the experiments were **medically justified**, the Nuremberg Military Tribunals condemned the experiments as "crimes against humanity"; 16 of the 23 physicians were found guilty and imprisoned, and 7 were sentenced to death. In the August 1947 v, the judges included a section called "Permissible Medical Experiments." This section became known as the Nuremberg Code and has formed the basis for ethics codes internationally.

1963: The Willowbrook Study

From 1963 to 1966, studies were carried out at Willowbrook State School, a New York institution for "mentally defective persons."

These studies were designed to gain an understanding of the natural history of infectious hepatitis and, subsequently, to test the effects of gamma globulin in preventing or ameliorating the disease.

The participants, all children, were deliberately infected with the hepatitis virus. Early participants were fed the stools of infected persons. Later, subjects received injections of more-purified virus preparations.

Researchers defended the deliberate injection of these children by noting that the majority would acquire the disease anyway while at Willowbrook, adding that perhaps it would be better for them to be infected under controlled research conditions.

During the course of these studies, Willowbrook closed its doors to new inmates, claiming overcrowded conditions. However, the hepatitis program was able to continue to admit new patients because it occupied its own space at the institution.

Thus, in some cases, parents found they were unable to admit their children to Willowbrook unless they agreed to their participation in the studies.

23.1 Foundations of Research Ethics

- **Nuremburg Code (1947): mandated voluntary consent for experimental studies of humans.**
- **Declaration of Helsinki (1964): written by the World Medical Association to provide guidelines for physicians conducting clinical trials,** and outline the principles of recruiting and involvement humans in researches that require interventions.
- **Belmont Report (1979): published by the U.S. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research to define key research principles and is a foundational document for the current U.S. federal policy for protecting human research participants (the Common Rule).**

** All patient protection regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule in the United States, must be strictly adhered to for observational as well as experimental studies.

23.2 Respect, Beneficence, and Justice

- **<u>Respect</u> for persons is a broad concept that emphasizes informed consent, voluntariness, and autonomy** (only an individual is authorized to decide whether to volunteer to participate in a research study)**.**

- **<u>Beneficence</u> means that the study should do good; <u>nonmaleficence</u> means that should do no harm.**

- **Distributive justice seeks to ensure that the benefits and burdens of research are equitable.**

❖ Questions that can be asked for community-based projects to complement key questions associated with individual-focused projects:

FIGURE 17-2 Sample Ethical Considerations for Individual- and Community-Based Research Projects

	Individual Participants	Community Participants
Respect	▪ What steps have been taken to protect individual rights? ▪ Has the risk of coercion in recruitment been considered and minimized? ▪ Is the informed consent process more than just signing a piece of paper? ▪ Do participants in sensitive studies have privacy? Will their participation be kept secret? ▪ Will data shared with the researchers be kept confidential? Will files be protected and not shared unless individually identifiable information is removed?	▪ What steps have been taken to ensure that a community's values are respected? ▪ Are appropriate community-based research methods being used? ▪ Have community representatives and a local oversight committee been consulted about the project?
Beneficence	▪ How will individuals benefit from participation? Free services, supplies, or medicines? Free health education? Gifts or money? Contribution to knowledge?	▪ How will a participating community benefit from the research project?
Nonmaleficence	▪ What steps have been taken to minimize physical, psychological, financial, social, and other risks to participants? ▪ Is counseling available for participants in sensitive studies? ▪ Is appropriate reimbursement for travel costs and other expenses being offered?	▪ What steps have been taken to ensure that a community is not burdened by research participation?
Justice	▪ What are the long-term benefits for individual participants? For example, will they gain increased knowledge about their health status? ▪ What will happen to participants after the study is completed? Will the results of the study be shared with them?	▪ What are the long-term benefits of participation to the community? ▪ Will the researchers have an ongoing relationship with the community?

❖ Questions that researchers should ask and answer about their own protocols prior to formal review by an ethics committee:

FIGURE 23-1 Eight Central Considerations ("8 Cs") in Research Ethics

Category	Examples of Questions to Ask
Contribution	• Why is the proposed project important? • How will individuals and/or communities benefit from this study?
Compensation	• Will individuals or communities that participate in the study be offered any form of inducement, reimbursement, or compensation? If so, what will be offered, and is it appropriate? Is the offer so high that it could be seen as coercive or so low that the study could be seen as exploitative? • Are the risks of participation minimal? • How will study-related injuries be handled? • Are the risks and benefits balanced?
Consent	• How will potential participants be informed about the study? • How will consent to participate be documented? • Will a test of comprehension of the informed consent statement be required? • If applicable, how will consent (and possibly assent) be acquired for children and other members of potentially vulnerable populations? • If applicable, will community meetings be held prior to beginning the study?
Confidentiality	• How will the privacy and confidentiality of participants and their personal information be maintained?
Community	• Why is research in the selected population important? • Is the source population appropriate for the goals of the research study? • Will the selection process be fair? • Will the sample size be adequate? • Are potentially vulnerable participants adequately protected? • Has the protocol been adapted to address the cultural expectations of the source population? • If applicable, has the community agreed to participate in this project?
Conflicts of interest	• Who is contributing to the project's finances and/or logistics? • Might potential conflicts of interest inhibit the ability of a researcher to conduct ethical and unbiased research?
Collaborators	• Are all members of the research team adequately trained to conduct ethical research? • What steps will be taken during data collection and analysis to ensure that the protocol and all ethical standards are adhered to by all members of the research team?
Committees	• Which research ethics committee(s) needs to review the project? • If applicable, what community organizations have been consulted about the proposed project?

23.3 Incentives (حوافز) and Coercion (إكراه)

- **The desire to thank participants must be balanced with the need for participation in any research project to be voluntary** (i.e., it is not permissible to exploit people's need for money, but you can give them a little money as a gift; specially if they need transportation or something…

- **Researchers have to be very transparent about what participants will gain from participation in a research study and what they will not gain.**

** Coercion could include social pressure or requests from authority figures that make it difficult for an individual not to agree to enroll in a study.

23.4 Informed Consent Statements

- **Informed Consent Statements provide essential information about research projects to potential research participants so that they can make a thoughtful decision about whether to enroll in a study.**

- **The statement must use clear, simple language that the reader understands.**

FIGURE 23-3 Content for the Informed Consent Statement

Content Area	Description
Research	A definition of "research" and a statement that the study involves research
Purpose	An explanation of the purpose and aims of the research process (except in the rare situations in which that interferes with the research goals)
Participants	A description of how and why certain individuals or communities were invited to participate in the research project and an estimate of the total number of individuals who will be recruited
Procedures	A description of the study procedures (including any physical exams, collection of biological specimens, randomization or blinding processes, interventions, or other procedures that are part of the study protocol) and the expected duration of the individual participant's involvement in the study
Benefits	A description of benefits to participants and/or to society, including a clear explanation of the compensation to be offered or a clear statement that the participant will receive no direct benefits
Risks	A description of the possible risks, discomforts, and costs associated with participation, a statement that involvement in the project may involve unforeseeable risks, and a description of how study-related injuries will be handled
Confidentiality	A description of the steps that will be taken to maintain confidentiality
Voluntariness	A statement that participation is voluntary and that the participant may withdraw from the study at any time with no penalty, along with a description for the process of withdrawing from the study
Contact information	Contact information for the researchers
Signature	Space for the participant's signature

23.5 Informed Consent Process

Informed consent is intended to be a process, not merely a piece of paper.

Acquiring a signature is not the end of the process; **the lines of communication between researchers and participants must remain open during and even after the data collection process;** because for any time the participants may decide not to continue and withdraw from the study!

23.6 Informed Consent Documentation

For most research studies, the expectation is that each study participant will sign a printed copy of the informed consent statement.

In a limited number of observational studies (i.e., studies without interventions like cross-sectional studies)**, the full process of acquiring and documenting individual informed consent may not be required.**

- A consent process that does not require a signature may be granted when:
- The responses cannot be linked to individuals.
- The survey instrument does not ask sensitive questions.
- The researchers will not physically examine individuals or collect biological specimens.
- The questionnaire is so short that describing the study would take longer than completing the questionnaire form.
- There are no foreseeable risks to participants.

23.7 Confidentiality and Privacy

- **Privacy: is the assurance that individuals get to choose what information they reveal about themselves.**
- **Confidentiality: is the protection of personal information provided to researchers.**

23.8 Sensitive Issues

Researchers asking questions about sensitive issues must decide ahead of time how to handle disclosures (such as disclosures of participation of illegal activities; drug or alcohol abuse, and sexual practices**).**

The research team can apply for a certificate of confidentiality that protects the identity of participants from being subject to court orders and other legal demands for information.

23.9 Cultural Considerations

A research protocol must be appropriate to the culture or cultures of the expected study participants.

It may be helpful to have a local advisory board facilitate communication between the community and the research team.

23.10 Vulnerable populations

Children and some adults with cognitive impairments may not be considered competent to make an informed decision.

Whenever possible, in addition to having the legal representative's consent, potential participants should assent to their own participation.

23.11 Ethics training and Certification

Research ethics committees usually require everyone who will be in direct contact with research participants and/or their personal data to complete formal research ethics training.

Responsible Conduct of Research (RCR) training programs may also spell out expectations and procedures for disclosing conflicts of interest, avoiding research misconduct, and exhibiting professionalism as researchers.

24.1 Ethics Committee Responsibilities

** **IRB** is a group responsible for protecting people who participate in research studies.

❖ **The three primary goals of Research Ethics Committees (RECs) often called Institutional Review Boards (IRBs), are to:**

1) **Protect the "human subjects" who will participate in research.**
2) **Protect researchers by preventing them from engaging in activities that could cause harm.**
3) **Legally protect the researcher's institution from the liability that could occur as a result of research activities.**

❖ **The major functions of ethics review boards are to:**

- **Review and revised research protocols.**
- **Approve or disapprove those protocols.**
- **Ensure that informed consent is documented (if required).**
- **Conduct continuing review of long-term research projects.**

** An **Institutional Animal Care and Use Committee** (IACUC) oversees research with animals and operates separately from an IRB

24.2 Ethics Committee Composition

Research ethics committees are usually composed of at least five members, preferably from diverse backgrounds, including both scientists and nonscientists (e.g., clergy/sheikh and lawyers)**.**

24.3 Application Materials

Some research ethics committees ask applicants to provide a narrative research statement that addresses a list of possible ethical concerns; others require the completion of dozens of pages of forms.

FIGURE 18-1 Examples of Information Requested and Examined by Ethics Review Committees

Category	Considerations
Participants	■ What is the anticipated composition and size of the study population? ■ Is the source population appropriate for the study question? ■ How will participants be recruited? Does the recruitment method raise any concerns about coercion? ■ What are the inclusion and exclusion criteria? Will the exclusion criteria screen out participants with a higher-than-typical risk of harm? Will the criteria generate a study population that is reasonably representative of the source population? (For example, if the study question applies to all adults, are there any restrictions on participation by reproductive-age women that are not directly related to safety?) ■ If applicable, are potentially vulnerable subjects protected?
Risks and benefits	■ Why is the study important and necessary? How will the proposed study benefit participants and/or their communities? ■ How will data be collected? Will existing data, documents, records, or specimens be used? Will individuals or groups be examined using surveys, interviews, focus groups, or other methods? Will interviews be audio- or video-recorded? Will noninvasive clinical measures be used? Will participants be asked to engage in exercise or tests of endurance, strength, or flexibility? What machines will be used to collect data, and will collection involve radiation exposure? Will blood, hair, nail clippings, sweat, saliva, sputum, skin cells, or other biological specimens be collected noninvasively? Will drugs or devices be tested? ■ What are the potential physical, psychological, financial, or other risks to participants? ■ Are the risks minimal (or at least minimized)? ■ Are the risks reasonable compared to the anticipated benefits?
Informed consent	■ Does the informed consent statement adhere to institutional guidelines? ■ How will informed consent be sought? ■ How will informed consent be documented? ■ Is any modification to the usual methods of documenting informed consent being requested? If so, is the request reasonable? (For example, is a waiver of a signed consent form being requested because the source population has a low literacy rate? Or is a request being made to have no documentation of consent because the existence of a form linking an individual to the study could harm the participant?)
Privacy and confidentiality	■ How will privacy and confidentiality be maintained? ■ What are the plans for the protection of computerized and noncomputerized data?
Safety monitoring	■ What constitutes an adverse event? How will such events be handled? ■ Does the informed consent statement clearly state how research participants can contact the research team and the ethics review board if they have concerns?
Conflicts of interest	■ How is the project being funded? ■ Do any financial or personal conflicts of interest need to be disclosed to participants and/or addressed in other ways?
Researcher training	■ Are the investigators prepared to conduct ethical research?
Documentation	■ Are copies of all recruitment materials attached? ■ Are copies of the questionnaire and other assessment tools attached? ■ Is a copy of the informed consent statement attached? ■ If applicable, are letters of approval from study sites and collaborating institutions attached? ■ If applicable, is a copy of the grant proposal attached? ■ Are copies of research ethics training certificates for all members of the research team attached?

24.4 Review Process

Important!

Once all application materials have been submitted to a research ethics committee, there are three possible next steps:

1) Exemption from review. Exemption can be granted—but does not have to be granted—only after the IRB professionals review a protocol and determine that it meets their criteria for exemption. When a researcher is considering transitioning from a practice-based inquiry (a clinician examining their patients) to an intentional research project (a clinician reviews patient records so that they can be presented as a case series at a professional Conference), the IRB should be consulted about what application materials are required. The decision about whether a practice-based project is exempt from review is up to the IRB, not the researcher.
** Exemption from review is not allowed for research focused on vulnerable populations.

2) Expedited review: is a determination by an IRB that a proposal requires review but a review by the full committee is not required. An expedited review may be possible when a minor change to a previously approved protocol is requested. Sometimes expedited review is also possible for new studies in which the risk to participants is no greater than what is encountered in ordinary daily life or, in the case of clinical work, during routine examinations or procedures.

3) Full review: is a determination by an IRB that the full committee must discuss a study protocol in order to ensure that the requirements for the protection of human subjects are met. Full review of a research proposal is usually required when an intervention will be tested in individuals or a community, data will be collected through interaction with individuals, identifiable private information will be collected, or other criteria for expedited review are not met.

** These decisions are made by research ethics committees (not the researchers themselves).

24.5 Review by Multiple Committees

At least three issues must be resolved prior to submission of a research proposal to multiple committees:

i.e., Some researchers may do research in multiple institutions; so they should meet each institution requirement.

- ✓ **The application documents that will be required.**
- ✓ **The wording of the informed consent statement.**
- ✓ **The order of review.** Sometimes, all the committees independently review the proposal at the same time. At other times, the reviews are conducted "domino" style, with the proposal being independently reviewed and approved by one committee, then passed to the next committee, and so on.

24.6 Ongoing Review

All ongoing research protocols must be re-reviewed annually (or more often, at the discretion of the ethics review committee). All adverse events must immediately be reported to the IRB. Any desired changes to recruiting materials, the informed consent statement, the questionnaire, or other study documents must receive approval prior to being implemented. At the end of a study, most committees require a final report to be submitted that at minimum states the number of participants, affirms that no adverse events occurred, and declares that the project is concluded.

24.7 Conflicts of Interest

When a financial or other relationship (personal relationships, board membership, or others) could bias the design, conduct, or reporting of the study, the potential Conflict of Interest (COI) must be disclosed.

The disclosure of a potential COI is not an admission of bias, but it is an important assurance of transparency.

24.8 Is Ethics Review Required?

Institutional approval provides a degree of legal protection to the researcher, and many research sponsors will not release grant or contract funds until a research plan has been approved by a research ethics committee.

**** The decision to exempt a project from review can be made only by the relevant ethics committees.**

**** Research protocols cannot be retroactively approved, so researchers must take the time to undergo a formal review prior to collecting any data or analyzing any data files.**

19

Chapter 16 : Population Sampling

An accessible and appropriate source of study participants for primary studies should be identified early in the research process.

16.1 Types of research Population

Target population **>** source population **>** sample population **>** study populations

Definitions :

→ Target population: the general population that the study seeks to understand and to which the results of the study should be applicable **(the broadest group).**

→ Source population: the subset of individuals from which a representative sample will be drawn.

→ Sample population: Individuals who are asked to participate.

→ Study population: Eligible participants who consent to participate in the study.

16.2 Target & Source Populations

→ Target population: the population to which the results of the study are intended to apply (for **generalisation**). Might be quit narrow

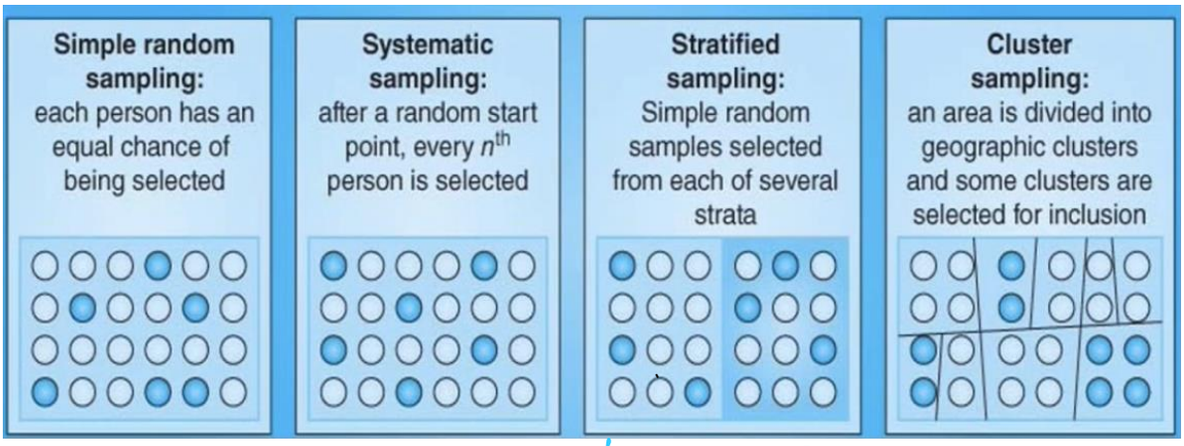
→ Source of population (sampling frame): a list of particular people from whom a sample population can be drawn.

→ Sample population: the members of the source population who are invited to participate in the study.

Ideally, probability based sampling is used to ensure that the sample population is representative of the source population.

Non random sample bias → that could occur. If each individual in the source population doesn't have an equal chance of being selected for the sample population. # researchers using a convenient sample must avoid the ascertainment bias that can occur if the convenience sample isn't Representative of the source population as a whole

Examples of types of probability sampling:



↳ Males & Females

✓ Non-probability-based convenience Population:
Selected based on ease of access, non-representative, which leads to non-random sampling bias. Participants characteristics should be compared with the broader community intended to represent, to avoid ascertainment bias (non-representative convenience population).

16.4 Study Populations

people who actually participate in the study.

Three things we should know:

- Few studies achieve a 100% participation rate among those who are invited to participate, but researches should aim for a high participation rate.
- Response rate: the percentage of people who participate in the research.
- Low response rate: rate leads to non-response bias.

16.5 Populations of cross-sectional Surveys

Avoid convenience samples that are not representative because they can't be generalized on the population, as the results are often used to make policy decisions.

FIGURE 16-3 Population Example for a Cross-Sectional Study

Study approach	Cross-sectional survey
Study question	What proportion of high school students in North County smoke cigarettes?
Data collection method	Participants will complete their own paper-based questionnaires.
Target population	Students in grades 9–12 in North County
Source population	All students enrolled in any of the 14 high schools in North County
Source population list	A list of the number of students in each homeroom provided by each high school
Sample population	Based on estimated sample size requirements, 20% of homerooms will be randomly sampled from the lists provided, and all students in these sampled homerooms will be asked to participate in the study.
Study population	Eligible individuals from the sample population who agree to participate
Confidentiality	No student names will ever be provided to researchers; surveys will be anonymous.

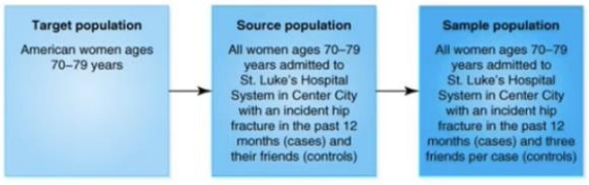
16.6 Populations for Case-Control Studies

Find the cases then identify an appropriate source of controls.

All cases must have the same disease, disability, or other health-related condition.

***The case definition should specify the inclusion and the exclusion criteria.**

FIGURE 16-4 Population Example for a Case-Control Study



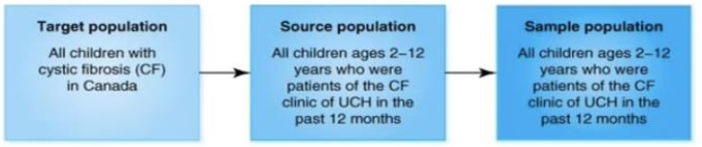
Study approach	Case-control study
Study question	What are the risk factors for hip fractures in adult women in the United States?
Data collection method	Participants will be interviewed in person or by telephone.
Target population	Women ages 70–79 living in the United States
Source population	All women ages 70–79 who were admitted to St. Luke's Hospital System in Center City with an incident hip fracture in the past 12 months
Source population list	A list of the hospital registration numbers for each inpatient female ages 70–79 at admission whose electronic medical record indicates a diagnosis of a hip fracture (ICD10 code S72) during the eligible 12-month period.
Sample population	All members of the source population will be asked to participate as cases, and each case will be asked to provide the names of three female friends in the same age range who might be able to serve as controls.
Study population	Eligible individuals from the sample population who agree to participate
Confidentiality	The hospital will provide the researcher with the names, addresses, and phone numbers of potential participants. Personally identifying information will not be included in the electronic file that contains questionnaire responses.

16.7 Populations for Cohort Studies

A longitudinal cohort study need a representative population (like a cross sectional study).

Prospective & retrospective cohort studies start by identifying an appropriate exposed population (in the same way that case-control studies start by identifying a source of cases).

FIGURE 16-5 Population Example for a Cohort Study



Study approach	Cohort study
Study question	What is the incidence rate for lung infections in children with cystic fibrosis?
Data collection method	Participants' parents will be asked to log all infections throughout the 2-year prospective study period, and these will be checked against the patients' medical records.
Target population	All children with cystic fibrosis in Canada
Source population	All children ages 2–12 years who were patients of the cystic fibrosis clinic of University Children's Hospital (UCH) in the past 12 months
Source population list	A list of all children ages 2–12 who were examined at the UCH cystic fibrosis clinic in the past 12 months
Sample population	The parents of all individuals in the source population will be asked if they will allow their children to participate in the study.
Study population	Eligible individuals from the sample population whose parents agree to let them participate
Confidentiality	All guidelines and regulations for the protection of patient information will be strictly adhered to, and only essential personnel will have access to patient records.

16.8 Populations for Experimental Studies

Be aware of special ethical requirements associated with interventional studies, **1st priority is safety**. (Not necessary that everyone who is recruited will continue the study, because sometimes there are decisions terminate participation of some participants based on safety issues).

Issues in experimental studies differ, it's all about matter of power not a matter of representation (having enough number in each group to test my hypothesis "hypothesis testing").

FIGURE 16-6 Population Example for an Experimental Study

Study approach	Experimental study
Study question	Does nutritional counseling during the first semester of college prevent weight gain?
Data collection method	Half of the participants will be assigned to meet weekly with a nutritionist during their first semester, and half will have no intervention. All participants will complete nutritional assessments during the first and last weeks of the fall and spring semesters of their first year at college.
Target population	First-year students at primarily residential colleges
Source population	All first-year students at East State College
Source population list	A list of all students enrolled in the mandatory first-year seminar class at East State College
Sample population	A randomly selected sample of students from the source population
Study population	Eligible individuals from the sample population who agree to participate
Confidentiality	Nutritional counseling and assessment sessions will be conducted in a private setting, and only essential personnel will have access to participants' records. Participation in the study will be voluntary, and professors teaching first-year seminars will not known which students have enrolled in the study.

Some experimental studies require participants to be exposed to risky substances or activities. In such studies, the risk of harm can be reduced by selecting an appropriate source population and defining strong inclusion and exclusion criteria. **For example:**

- studies that involve exercise must target potential participants likely to be healthy enough to engage in physical activity.
- Studies of new drugs for advanced forms of cancer are often open only to extremely ill patients for whom standard therapies have not been effective.

16.9 Vulnerable Populations (when we select our participants for our research)

- Vulnerable population in health research include Young children, people in prisons and people with limited ability to make an informed autonomous decision about participation (need special procedure to participate).
- These populations should not be selected for studies that don't require their participation and even when they are involved, **pay attention to ethics of research with vulnerable groups.**

16.10 Community Involvement

Community involvement:

- Improves response rate.
- Improves cultural competence of the research team.
- Ensures that outcomes are valuable to the community.

Community – based participatory research (CBPR): is based on partnerships in which community members identify research priorities and are involved in every stage of the research process.

*Used a lot in health research.

20

Chapter 17 : Sample Size Estimation

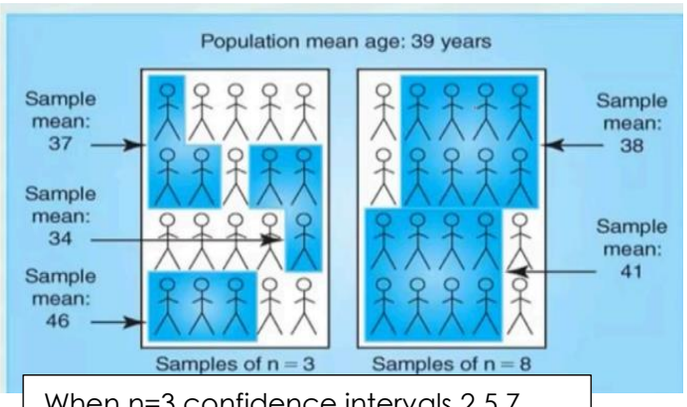
17.1 Importance of sample Size

- Based on estimation.
- Recruiting too many participants wastes resources.
- Recruiting too few participants makes the study invalid (so there is a minimal number of participants based on the research **question** and **design** and **how you are going to answer such question based on statistical analysis**).

If you can answer your research questions with 100 participants then why you go to 500? **The goal is to recruit just the right number of participants, not too many and not too few.**

17.2 Sample Size & Certainty Levels

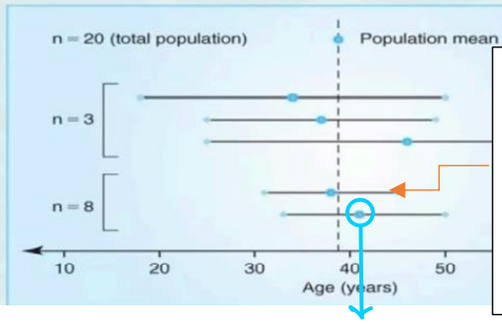
- Larger sample size produces narrower confidence intervals (an estimate of how close to the population value a sample of a particular size is expected to be) for statistical measures.
- Larger sample sizes yield more statistically significant results.



When n=3 confidence intervals 2,5,7
When n=8 confidence intervals 1,2

39 − 37 = 2
39 − 34 = 5
39 − 38 = 1 ...

Figure 17-2: **Larger Samples from a Population Have a Narrower 95% CI Than Smaller Samples**



This line represent 95% confidence interval samples that were taken and closer to population mean .

It's important to have reading btw the mean of your Variables in population as close as possible to mean of origin population

the mean of my sampled population

If you include all the population in the analysis there will be no need for confidence intervals. You have the react value

17.3 Sample Size Estimation

- Sample size calculators (estimators) estimate the number of participants necessary for a study based on guesses about the likely results of the study.
- When the level of certainty about inputs is low, it's better to err on the side of a larger sample size((**waste more money much better than losing the study validity**), that's why after the calculation we always add a little more to avoid drop outs.)

FIGURE 17-3 Examples of Sample Size Calculation

Characteristic	Cross-Sectional Survey	Case-Control Study	Cohort Study	Experimental Study
Study question	What proportion of the population has the exposure or disease?	Are cases more likely than controls to have the exposure?	Are exposed people more likely than unexposed people to develop the outcome?	Are exposed people more likely than unexposed people to have a favorable outcome?
Population size	5000	—	—	—
Anticipated percentage with exposure or disease	15%	—	—	—
Confidence for anticipated exposure percentage	±3%	—	—	—
Ratio of controls to cases	—	2	—	—
Ratio of unexposed to exposed	—	—	1	1
Anticipated percentage of controls exposed	—	25%	—	—
Anticipated percentage of unexposed with disease or outcome	—	—	10%	70%
OR worth detecting	—	1.5	—	—
RR worth detecting	—	—	1.3	1.25
Confidence level (1 – α)	95%	95%	95%	95%
Power (1 – β)	—	80%	80%	80%
Estimated sample size	~500	~350 cases and 700 controls	~1850 exposed and 1850 unexposed	~90 exposed and 90 unexposed

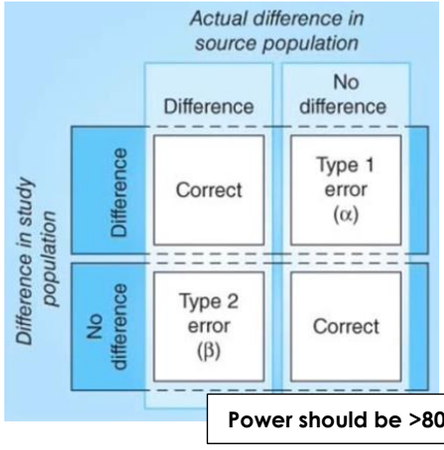
FIGURE 17-4 Sample Size Estimates for a Case-Control Study

Situation	Ratio of Controls to Cases	Anticipated Percentage of Controls Exposed	Anticipated Percentage of Cases Exposed	Odds Ratio (OR) Worth Detecting	Estimated Sample Size Required	Estimated Number of Cases Required	Estimated Number of Controls Required
Base case (Figure 17-3)	2	25%	33%	1.5	1050	350	700
1:1 ratio	1	25%	33%	1.5	950	475	475
10:1 ratio	10	25%	33%	1.5	2750	250	2500
Lower % exposed	2	10%	14%	1.5	2025	675	1350
Higher % exposed	2	30%	39%	1.5	976	325	650
Lower OR	2	25%	29%	1.2	5400	1800	3600
Higher OR	2	25%	40%	2.0	345	115	230

17.4 Power Estimation

- Type 1 ERROR(α): the study says there is a significant difference but it doesn't exist in actual population. Usually =5% and corresponds to 95 CI (some studies make it 1% with larger sample size) and is written as P-value of p =.05. (In significant studies P- Value should be < than .05)
- Type 2 ERROR(β): there is a difference in actual population, but the study says there is no difference. 20% is acceptable level (power=1-β =80%)

- Power is the ability of a statistical test to detect significant differences between subgroups of a population when differences **really** do exist.
- Studies with more participants have more power.



Power should be >80%

FIGURE 17-6 Examples of Power Calculation

Characteristic	Cross-Sectional Survey	Case-Control Study	Cohort Study	Experimental Study
Number of exposed	100	—	2500	70
Number of unexposed	250	—	1500	70
Number of cases	—	250	—	—
Number of controls	—	490	—	—
Percentage of exposed with disease or outcome	40%	—	13%	85.7%
Percentage of unexposed with disease or outcome	26%	—	9%	64.3%
Percentage of cases with exposure	—	32%	—	—
Percentage of controls with exposure	—	25.5%	—	—
Confidence level (1 – α)	95%	95%	95%	95%
Estimated power (1 –β)	~70%	~45%	~97%	~80%

17.5 Refining the Study Approach

- Recruit larger sample size than the estimated.
- The approach must be adjusted (increase source population) if the number of likely participants is much lower than the estimated number of people required for adequate power.

The effect size: the difference between the exposed versus not exposed or 2 groups (how much are the effects of the variables that we are studying on the percentages of the attributes that we are detecting on the outcomes).

(تلمّح فجر الأجر يهُن ظلام التكليف) ابن الجوزي.

18.1 Questionnaire Design Overview:

Questionnaire (or **survey instrument**): is a series of questions used as a tool for systematically gathering data from study participants.

- A good questionnaire is carefully crafted for a specific purpose
- Most research studies must design new data collection instruments, but for some research topics, validated question banks are available and questions can be selected from them.

18.2 Questionnaire Content:

- Start by compiling a list of the topics the survey instrument must cover. (These usually come from your main consent)

➔ **For example**: if you are collecting data about prevalence of smoking among students, you think of the question that makes relationships with this phenomenon that you're interesting in. (is age important for you? Is the income of the family important? Is the living place important? Is educational level important? Etc...)

➔ A source for these types of questions (your content) usually comes from the literature and based on the topic of your research.

- Include questions that confirm eligibility to participate.

➔ You should be asking questions that confirm that such participant is eligible to be included or not.

For example: if you're aged less than 17, please DON'T answer this questionnaire.

❖ **The first set of questionnaire items typically enables the researcher to confirm that participants meet the eligibility criteria for the study**. For example, if only currently registered students are eligible to participate in a university-based cross-sectional survey, one of the first questions should be about school enrolment status. Participants who report that they are not students must be excluded from analysis based on that answer

These are the major questions that usually could be asked in a questionnaire:

Systems thinking: is the process of identifying the underlying causes of complex problems so that sustainable solutions can be developed and implemented.

➔ May include potential cofounders: factors that might influence the relationships between key exposures and outcomes. Ex: smokers consume more alcohol so higher liver disease rate, so if you're studying the relation between smoking and liver disease alcohol is a cofounder variable.

18.3 Types of questions:

- **Closed ended questions (close-ended questions) allow a limited number of possible answers:**
    - Date/time
    - Numeric
    - **Categorical**: **ordinal** (ranked) or **nominal** (unordered)
- **Open-ended**, or **free-response**: questions allow participants to explain their answers at length.

❖ Most questionnaires questions are usually closed-ended.

❖ Closed-ended questions are usually easier to statistically analyse than open-ended questions.

Examples of types of questions →

✓ Yes or no questions are called dichotomous or binary.

✓ Nominal (unordered) questions there is no order in the answers, For example: what's your type of film? Whatever you answer, action, comedy, or documentary, it's still a preference and there is no advantage of selecting any over the other.

✓ Example on ranked questions: educational level.

What's your highest educational level?

Less than high school -> 1 (we give them one point)

High school -> 2

College without degree -> 3

College -> 4

Postgraduate -> 5

So, now there's ranking.

✓ In paired questions, you give them a choice of two, and you give them an alternative as well.

FIGURE 18-3 Examples of Types of Questions

Type	Sample Question	Sample Response Option for the Sample Question
Date	What is your birth date?	__ __ - __ __ - __ __ __ __ m m - d d - y y y y
Numeric	What is your height without shoes (rounded to the nearest half inch)?	__ __ . __ inches
Yes/no	During your lifetime, have you smoked more than 100 cigarettes?	❑ Yes ❑ No
Categorical/ multiple-choice: nominal (no rank)	What is your sex? What is your favorite type of film?	❑ Female ❑ Male ❑ Action/drama ❑ Comedy/musical ❑ Documentary ❑ Other: _____
Categorical/ multiple-choice: ordinal (ranked)	What is the highest level of education you have completed?	❑ Less than high school ❑ High school ❑ Some college but no degree ❑ College/university degree or advanced degree
	How much do you agree with this statement?: "No matter how much I exercise, I will not be able to lose weight."	❑ Strongly disagree ❑ Disagree ❑ Neutral ❑ Agree ❑ Strongly agree
	On a scale of 1 to 5, with 1 meaning poor and 5 meaning excellent, how would you rate your hearing (without the use of a hearing aid)?	Poor_____Excellent ❑ 1 ❑ 2 ❑ 3 ❑ 4 ❑ 5
Paired-comparisons	Do you prefer to drink coffee or tea?	❑ I prefer coffee ❑ I prefer tea ❑ I like coffee and tea equally ❑ I do not drink coffee or tea
Rank-ordering	List the following four political issues in order from most important to you (1) to least important to you (4): crime/safety, environment/energy, foreign policy/defense, taxes/revenue	Number from 1 (most important) to 4 (least important): ___ Crime/safety ___ Environment/energy ___ Foreign policy/defense ___ Taxes/revenue
Open-ended/ free-response	What is your biggest personal health concern at present?	_____ _____

18.4 Anonymity:

- When anonymity is important, avoid asking questions that could allow the participant's identity to be determined based on his response.
- For many types of studies, there is no need to collect names, contact information, or other identifiable information.

➔ You should not include any question that could lead to the breakdown of the ID of the participants.

18.5 Types of responses:

- Provide clear instructions about the acceptable types of answers for each question.
- Add response for "not applicable" when relevant.

FIGURE 18-4 Examples of Five-Point Responses for Ranked Questions

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Dissatisfied	Somewhat dissatisfied	Neutral	Somewhat satisfied	Satisfied
Very negative	Somewhat negative	Neither negative nor positive	Somewhat positive	Very positive
Poor	Fair	Good	Very good	Excellent
None	Few	Some	Many	Very many
Never	Rarely	Sometimes	Often	Always
Not important	Slightly important	Somewhat important	Very important	Extremely important

➔ This scale is called (**Likert scale)** which presents ordered responses to a questionnaire item that asks participants to rank preferences numerically, such as by using a scale for which 1 indicates strong disagreement and 5 indicates strong agreement.

➔ There's five points possibility, when do we give 1, 2, 3, 4 or 5? It depends...

- Some scales the higher the number the better the answer,
- Some scales the lower the number the better the answer,
- like this question: did you feel tired during the last week? Never here is better than always, those are called negative questions, and need to be recoded when we deal with the data that we're collecting.

18.6 Wording of questions:

- Check each question and all the response items for clarity.

FIGURE 18-5 Problems to Avoid

Problem	Example	Problem with the Example
Big words/ jargon	Have you ever had a myocardial infarction?	Participants may not know that a "myocardial infarction" is a technical term for a heart attack.
Undefined abbreviations	Have you ever been told that you have BPH?	Participants may not know that BPH is short for benign prostatic hypertrophy or that BPH means an enlarged prostate.
Ambiguous meanings	What kind of house do you live in?	Without seeing a list of appropriate responses, it is not clear if the answer should be "an apartment," "a rental," "a split-level duplex," or "a single-family home."
Vagueness	Do you exercise regularly?	"Regularly" is not defined. A person who exercises most days of each week might assume that "regularly" means daily and say "no." Another person who exercises once a month may consider that regular. It would be better to ask "In a typical week, how many days do you exercise for at least 30 minutes?"
Double negatives	I did not find this visit with my doctor to be unpleasant. ❑ Disagree ❑ Neutral ❑ Agree	The wording of this question makes it hard to figure out whether a person who was satisfied with a visit should agree or disagree.
Faulty assumptions	Do your gums bleed during regular dental cleanings? ❑ Yes ❑ No	The question assumes that everyone has routine dental cleanings. If "I do not visit the dentist" is not an answer option, a person who does not have dental cleanings is forced to answer no.

FIGURE 18-5 (continued)

Problem	Example	Problem with the Example
Leading answers	What is your impression about the quality of services provided by Center City Hospital? ❑ Fair ❑ Good ❑ Great ❑ Excellent	This question's response options clearly are intended to lead to a positive response; there is no "poor" option.
Answers with a poor scale	How many hours a week do you watch television? ❑ 0 ❑ 1–3 ❑ 4–7 ❑ 8 or more	Even though most people watch more than 1 hour of television daily, which would put them in the "8 or more" response category, they may not want to choose an "extreme" answer. Their inaccurate responses will lead to a false report. Alternatively, these response options may cause respondents to misread the question as how many hours a *day* they watch television.
Lack of specificity	What is your income?	It is not clear if income refers to earnings per hour, week, month, or year, or whether it refers to pre- or post-tax income.
Missing answer options	What color are your eyes? ❑ Brown ❑ Blue	Many possible eye colors are missing.
Overlapping answer options	In a typical week, how many days do you eat fish? ❑ 0 ❑ 1–3 ❑ 3–5 ❑ 5–7	Participants who eat fish 3 days a week or 5 days a week will not know which response to select.

18.7 Order for questions:

- Start with easy questions before moving to more difficult or sensitive questions.
- Avoid **habituation**: once responders have given the same answer to so many questions (agree, agree and agree) that they continue to reply with the same response (to avoid mix up questions).

18.8 Layout and formatting:

- Provide instructions
- Use readable and large fonts and include adequate white space.

FIGURE 18-6 Example of a Self-Response Questionnaire

Basic Information

1. What is today's date? __ __ - __ __ - __ __ __ __ (m m d d y y y y)
2. What is your date of birth? __ __ - __ __ - __ __ __ __ (m m d d y y y y)
3. What is your sex? ☐ Female ☐ Male ☐ _____

Health History (*Check one answer box for each question.*)

4. Have you ever been diagnosed with breast cancer? ☐ Yes ☐ No → *If **No**, then skip to Question 8.*
5. Have you had a mastectomy (either partial or complete)? ☐ Yes ☐ No

FIGURE 18-7 Example of a Telephone Interview Script

Fill in today's date. __ __ - __ __ - __ __ __ __ (m m d d y y y y)

Read: *Thank you for agreeing to participate in this health study.*
I'm going to start by asking you some basic questions.

1. What is your date of birth? __ __ - __ __ - __ __ __ __ (m m d d y y y y) ☐ Refused to answer
2. What is your sex: female or male? ☐ Female ☐ Male ☐ Other: _____

Read: *Now I'm going to ask you a few questions about your medical history.*

3. Have you ever been diagnosed with breast cancer? ☐ Yes ☐ No → *If **No**, then skip to Question 7.*
4. Have you had a mastectomy (either partial or complete)? ☐ Yes ☐ No ☐ Refused to answer

➔ **A filter question or contingency question** is one that determines whether the respondent is eligible to answer a subsequent question or set of questions. For example, participants who indicate that they have never used tobacco products can be prompted to skip a series of questions about smoking habits. For paper-based forms, instructions for skips must be carefully described in words.

18.9 Reliability & validity: مهم

- Reliability/precision ثبات الأداة : consistent answers are given to similar questions.

➔ Look for internal consistency (Cronbach's alpha, KR-20) and test-retest reliability.
➔ Reliable means consistent, to what extent is it precise.
➔ Consistency means if you administer this questionnaire to the same individual more than one time, it will give you similar responses

❖ One aspect of reliability is **internal consistency**, which is present when the items in a survey instrument measure various aspects of the same concept. Some survey instruments ask the same question several different ways, or ask a series of similar questions, in order to confirm the stability of participants' responses.
❖ For example, a questionnaire **might include two questions that are opposites of one another**, like "I enjoy eating most fruits" and "In general, I do not like to eat fruit." The expectation is that all respondents who say the first item is true will say that the second item is false, and vice versa. If a very high proportion of respondents' answers meet this expectation, then that is evidence that the responses are reliable.
❖ internal consistency can be confirmed with tests of intercorrelation such as **Cronbach's alpha and the KR20**.
❖ **Intercorrelation** is present when two or more related items in a survey instrument measure various aspects of the same concept.

- **Cronbach's alpha:** is a measure of internal consistency that is used with variables that have ordered responses.
- **The Kuder-Richardson Formula 20 (KR-20):** is a measure of internal consistency that is used with binary variables

➔ Both of these statistics are expressed as a number between 0 and 1**. Scores near 1 indicate an assessment tool with minimal random error and high reliability**. Agreement between two or more evaluations is another facet of reliability

- **Test–retest reliability** is demonstrated when people who take a baseline assessment and then retake the test later have about the same scores each time they are tested.
- **Interobserver agreement, or inter-rater agreement** describes the degree of concordance among independent raters assessing the same study participants.
- **The kappa statistic, also called Cohen's kappa** and represented with the Greek letter κ, determines whether two assessors who evaluated the same study participants agreed more often than is expected by chance. **a valid study will have a value of kappa that is close to 1.**

$$\text{Kappa coefficient} = \kappa = \frac{\text{Observed agreement} - \text{Expected agreement}}{1 - \text{Expected agreement}}$$

"Several approaches are used to evaluate the accuracy of assessment tools that rely on self-reporting, such as psychometric tests and surveys about attitudes and perceptions. These tools are often considered to be proxy measures for an underlying theoretical construct that cannot be directly measured. For example**, happiness and intelligence cannot be directly measured with physical or chemical tests**, but survey instruments can be designed to measure them indirectly. Some researchers use **the word concept** to describe a theory informed by observations and use the term construct to describe a theory informed from more complex abstractions. For example, there is no particular threshold at which an object becomes "heavy" or a person becomes "rich," **because those definitions will vary for different individuals**, but measurements of weight and income can guide the evaluation of these concepts. In contrast, notions about constructs like "trust" and "leadership" are more difficult to quantify. Other researchers consider a concept to be a general abstraction and a construct to be a multidimensional concept that has been carefully defined and crafted for research purposes. In practice, **the terms "concept" and "construct" are often used interchangeably**"

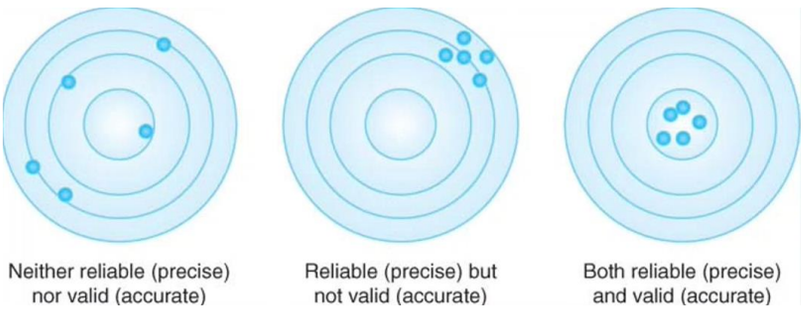
- Validity/accuracy صدق الأداة : the response or measurements are correct

Types of validities:

- **Content (logical) validity:** is present when subject-matter experts agree that the questionnaire captures the most relevant details about the study domain. Usually reviewed by an expert panel.
- **Face validity:** is present when content experts and users agree that a survey instrument will be easy for study participants to understand and correctly complete.
- **Construct validity:** the test measures the theoretical construct it intended to assess.
- **Convergent validity**: is present when two items that the underlying theory says should be related are shown to be correlated
- **Discriminant validity**: is present when two items that the construct says should not be related are shown not to be associated
- **Criterion (concrete) validity:** uses an established test or outcome as a standard (or criterion) for confirming the utility of a new test that examines a similar theoretical construct
  Example on criterion validity: a new test of intelligence can be validated against standard IQ tests

- **Concurrent validity:** is evaluated when participants in a pilot study complete both the existing and new tests and the correlation between the test results is calculated. A strong correlation between the tests is evidence that the new test is valid
- **Predictive validity:** is appraised when the new test is correlated with subsequent measures of performance in related domains
- **Factor analysis:** uses measured variables to model a latent variable that represents a construct that cannot be directly measured with one question but appears to have a causal relationship with a set of measured variables.

"Some statistical methods can provide information about which items in an assessment tool might be redundant or unnecessary and therefore can be removed without compromising the validity of the survey instrument. For example, **principal component analysis (PCA)** creates one or more index variables (called components) from a larger set of measured variables. The index variable is generated from the linear combination of measured variables, so it is a weighted average of the contributing variables. The PCA process determines the optimal number of components, the best measured variables to combine, and the best weights to use for the calculation"



18.10 Commercial research tools:

- Already-validated surveys instruments may be useful for incorporation into a new study.
- These may need to be purchased.

➔ The scale should be reliable, consistent, and accurate.

several widely used and validated tests are available to researchers, such as:

- The Beck Depression Inventory and the General Health Questionnaire (GHQ), which assesses psychological status.
- The Mini-Mental State Examination (MMSE), which evaluates cognitive function.
- The SF-36 and SF-12, which both measure health-related quality of life, a multidimensional construct that captures an individual's perceived physical, mental, emotional, and social well-being and the perceived impact of health status on the quality of daily life.

18.11 Translation:

Translation of the survey instrument into one or more additional languages may be necessary if the source population contains speakers of more than one language. Researchers using multiple languages must be certain that the translated version expresses the same meaning as the original survey. Accuracy may require the rephrasing of whole sentences, not just direct word-for-word translations.

There are two ways to ensure that the correct meaning is being conveyed:

1. **Back translation (double translation): in which one person translates the questionnaire from the original language to a new language and then a second person translates the survey instrument in the new language back into the original language.** ensures that the meanings of the original survey are maintained in the translated version.
2. A second approach is **to have two translators independently translate the survey instrument from the original to the new language**. The two translations are compared, and a consensus process is used to decide which words and phrases best convey the precise meaning and complexity of the original questionnaire.

18.12 Pilot testing:

-A **pilot test (or pretest):** is a small-scale preliminary study conducted to evaluate the feasibility of a full-scale research project**.**

➔ allows the researcher to correct problems with the survey instrument prior to data collection from participants.

A pilot test of a questionnaire is helpful for checking, among other issues:

1. The wording and clarity of the questions
2. The order of the questions
3. The ability and willingness of participants to answer the questions
4. The responses given, and whether the responses match the intended
5. The amount of time it takes to complete the survey

The researcher should ask several volunteers to help with the pilot test. These volunteers should be from the target population and meet the eligibility criteria for the study (in terms of age, exposure and disease status, and other key factors), but they should not be members of the sample population

22

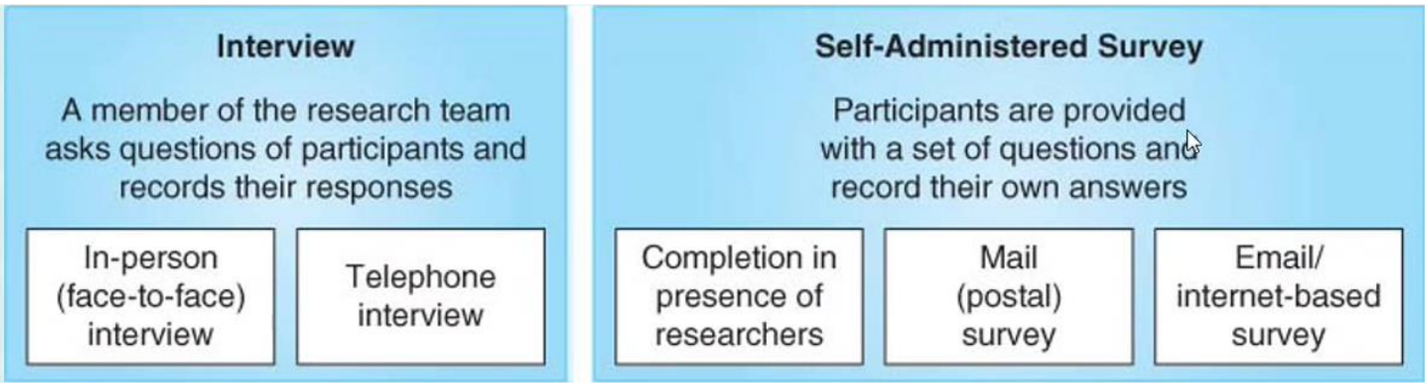
Chapter ~~19~~: Surveys and Interviews

19.1 Interviews versus Self-Administered Surveys:

Interviews are very popular and valid method of data collection, and they could be used by researcher to improve response rate and improve accuracy of data.

Most primary studies collect data from individual participants using an interview method or self-administered questionnaire. **Self-reported surveys are usually the least costly and least time-consuming way to gather information.** However, interviews may allow for more detailed information to be gathered and can be accompanied by laboratory and other tests.

- ❖ The first decision to make about collection of quantitative data is whether to have a member of the research team interview participants or to have participants record their own answers.
- ❖ Interviews may promote accuracy and completeness buy may be expensive because of personnel costs.
- ❖ Self-administered surveys may yield a higher sample size, be more cost-effective, and be preferable for sensitive questions.



An Interview: is the process of a researcher verbally asking a participant questions and recording that person's responses

A self-administered survey: uses a questionnaire form that participants complete by themselves, using either a paper-and-pencil version or an online version of the survey instrument

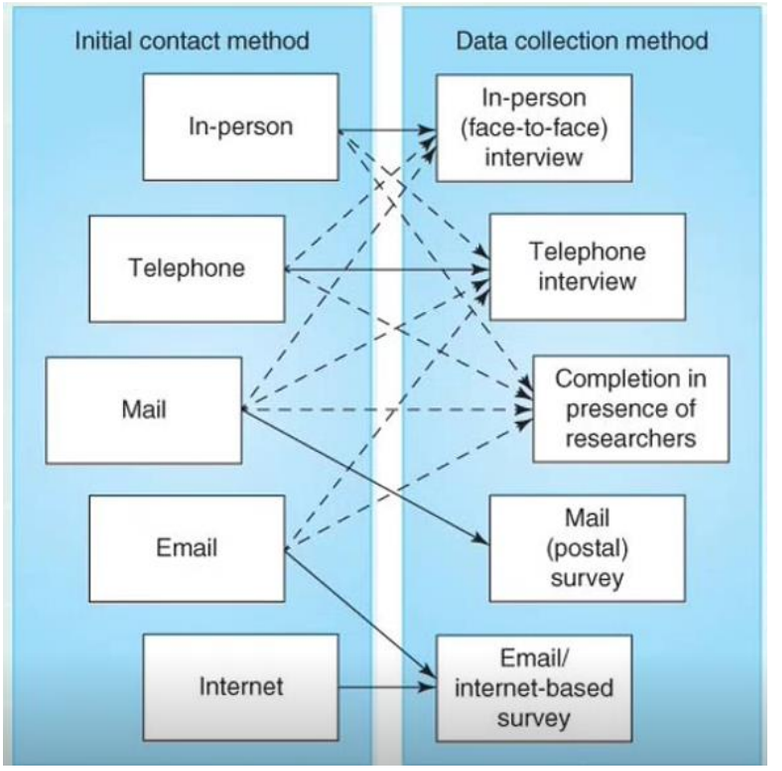
Interview	Self-administered surveys
Advantages: ❖ trained interviewers ensure the accuracy and completeness of each questionnaire while recording the responses. **Disadvantage:** ❖ may be expensive because of personnel costs	**Advantages:** ❖ allows for the cost-effective collection of data from a large number of participants. ❖ May be the best way to get honest answers to sensitive questions (like questions about mental or sexual health). ❖ Can be completed at specific study site, such as a workplace or school or hospital, or they can be delivered by mail or the internet.

19.2 Recruiting Methods:

- The goal is to maximize the participation rate among sampled individuals.
- The recruitment methods are often paired with data collection method (such as using online methods to recruit for an Internet-based survey).
- Provide multiple opportunities to participate, such as follow-up mailing (some participants don't respond from the first time, so you should follow with them).
- Consider appropriate incentives. (Incentives such as small gifts or the opportunity to be entered into a drawing to win a prize)
- Participation rates will likely be higher if recruits understand the importance and value of the research project.
- several hundred calls made by **random-digit dialing**—calls to a computer-generated list of unscreened telephone numbers—may yield only a few people willing to participate in a study

Examples of methods for contacting members of the sample population:

- If the plan is to **interview people in person**, the best recruiting method may be to visit potential recruits at work, at school, at home, at a public venue, or at another appropriate location. Alternatively, if the contact information for sampled individuals is available, which would be true if recruiting patients from a collaborating clinic or recruiting employees from a cooperating corporation, then interviews could be set up by sending a letter or an email of invitation and then following up with calls to all of the sampled individuals
- If the plan is to **interview by telephone**, the participation rate will be higher if a letter invitation is sent first.
- If the plan is to collect data **via the internet**, the most effective method is contacting potential participants via email or website.



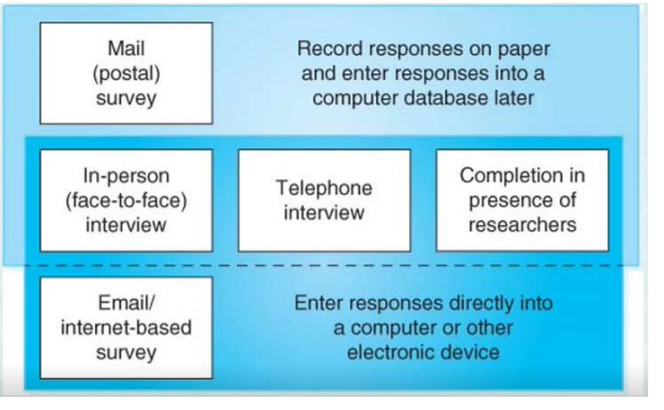
19.3 Data Recording Methods:

A decision must also be made about how responses will be recorded and when they will be entered into a computer database. There are two basic options.:

1. One is to record the responses on **paper** and to enter or scan them into a computer database later.
2. The other is to have interviewers or participants enter responses **directly into a database**.

❖ Paper-based surveys must later be scanned or typed into computers**; this can be expensive and time-consuming**, but they also **allow a large number of people to take the survey at one time** and place. (Also, error increases by this method, like missing data or data entry errors)

❖ Computer-assisted surveys input survey responses directly into a computer and can have built-in checks and "skips", but **not all population are comfortable with computers**. (**This method eliminates the need for later data entry, also missing of data is very minimal**)



19.4 Training Interviewers:

- **Uniformity** in data collection procedures is important.
- Interviewers need access to comprehensive handbooks and training sessions when they can practice their skills.
- The interview process should be the same for all participants in a study. Uniformity is easiest to accomplish when all interviewers are provided with the tools they need to follow a standardized set of procedures.

- **Interviewer bias** is a form of information bias that occurs when interviewers systematically question cases and controls or exposed and unexposed members of a study population differently, such as probing individuals they believe to have the disease or exposure of interest for more information but not doing the same probing for participants they believe to be unexposed controls.
- interviewer handbook that provides information about the purpose of the study, details about interview logistics

➔ The training and handbook should:

- Explain the interview process step-by-step
- Specify exactly how to ask questions and record responses
- Identify any prompts or follow-up questions that the interviewer must use or is allowed to use
- Emphasize any restrictions against asking for clarification about particular items
- Provide checklists for handling problems that might arise during an interview, such as interruptions

FIGURE 19-4 Characteristics of Well-Trained Interviewers

Characteristic	Actions That Demonstrate the Characteristic
Respectful	• Communicates pleasantly and professionally with all study participants and members of the research team • Has practiced interviewing enough to be comfortable with both the script and the interview process • Asks supervisors for assistance when it is needed
Organized	• Begins each scheduled interview session on time • Has all necessary materials on hand prior to the start of each interview session • Maintains meticulous records and completes all files and paperwork promptly
Considerate	• Dresses and grooms appropriately for in-person interviews • Is alert to modifiable conditions that may make interviewees uncomfortable, such as loud background noises or dim lighting • Allows adequate time for participants to respond to each question
Articulate	• Speaks at an appropriate pace and volume • Enunciates clearly • Uses an appropriate tone of voice (and, for in-person interviews, appropriate facial expressions and gestures) • Rereads questions and/or the list of closed-ended responses when a participant does not understand the question or the acceptable responses

FIGURE 19-4 Characteristics of Well-Trained Interviewers (continued)

Characteristic	Actions That Demonstrate the Characteristic
Consistent	• Reads the script exactly as it is written • Probes for answers only when the script indicates that probing is approved • Does not provide explanations for any question unless an explanation is provided in the script or approved in the interviewer handbook
Impartial	• Avoids verbal and nonverbal expressions of approval or disapproval • Does not express personal opinions • Avoids leading interviewees toward a particular answer (for example, by placing special emphasis on particular words in a question or by probing until receiving a particular desired response)
Honest	• Does not fabricate or falsify reports • Records responses to open-ended questions verbatim, without rephrasing, paraphrasing, "correcting," or interpreting them
Careful	• Completes all steps of the interview process in the correct order, as prescribed by the interviewer handbook • Documents informed consent prior to conducting an interview • Does not skip any component of the interview • Completes all response forms correctly

*Chapter 23 is NOT included in the course

Chapter ~~20~~ 24: Additional Assessments

20.1 Supplementing Self-Reported Information:

Self-reports, such as those made during interviews and the completion of questionnaires, are essential data sources, but they have significant limitations. Respondents may not tell the truth, either because they do not accurately remember the answers or because they want to provide answers that are thought to be correct.

❖ Laboratory tests and other objective measures can be used to supplement and validate self-reported data. These types of scientific data are usually collected in person by a member of the research team so that the measurements can be collected according to precise protocols

20.2 Anthropometric Measures:

Anthropometry is the measurement of the human body.

- **Examples include**: height, weight, waist circumference, Hip circumference, Mid-upper-arm circumference (MUAC) and skinfold measurement that estimate the body fat percentage.
- Ensure privacy for participants is being measured.
- It is often best for **two members** of the research team to be present when measurements are being taken.
- When a child is being measured, it may be necessary for a parent or guardian to be present with the child.

➔ **Standard methods** should be used to take all anthropometric measurements. Any tools used for the measurements should be carefully calibrated **to ensure accuracy and reliability**.

20.3 Vital Signs:

Vital signs: are physiological measurements that provide clinical data about an individual's essential body functions. Most basic vital signs can be quantified accurately after minimal instruction.

- **Examples include:** body temperature, blood pressure, pulse (heart rate), and respiratory rate.
  (A thermometer measures body temperature. A manual or electronic sphygmomanometer (a blood pressure cuff) measures systolic and diastolic blood pressure. Measuring resting pulse and respiratory rates does not require any instruments other than timekeeping devices)
- The research protocol should state exactly how each measurement should be taken.

➔ **Standardization increases the precision and validity of the measurements**. Additionally, tests of inter-rater reliability (discussed later) can be used to confirm that all assessors generate similar or identical results when they measure the same person.

20.4 Clinical examination:

Examples include:

- Heart sounds
- Breath sounds and other respiratory functions
- Bowel sounds and the condition of the abdomen
- The range of motion (ROM) and the condition of the joints
- The condition of the skin, hair, and nails
- The health of the eyes, ears, nose, and mouth
- Mental status
- The ability to conduct activities of daily living
- Other signs of health or disease

➔ **An assessment form can guide the procedures that will be used for all participants.**

Assessment form includes each component of examination, including exact procedures to be used and the specific diagnostic criteria for each item on the assessment form, as well as the order in which these elements should be examined. Care should be taken to ensure the comfort, privacy, and safety of each person being assessed.

20.5 Tests of Physiological Function:

Tests of physiological function can provide helpful data about health status.

- **Examples include:** spirometry for lung function, electrocardiography for heart function, electroencephalography for brain function, and audiometry for hearing acuity.
- Consider the costs of these test
  - ➔ Because of cost considerations, secondary analyses of existing medical records may be the best option for researchers whose study questions require the use of expensive equipment

20.6 Laboratory Analysis of Biological Specimens:

- **Examples include:** immunologic, genetic, and other tests for blood, urine, saliva, and other body fluids.
- Decide ahead of time whether results will be shared with participants.
  - ➔ Some immunologic, genetic, and other studies require the collection of new body fluids or tissue biopsies, either as part of routine clinical practice or specifically for the purposes of the research project. Before new specimens are collected, a research ethics committee must verify that the potential physical risks to participants caused by the collection of the sample will be minimized. Some studies may be able to make use of existing specimen banks. These samples may be fully anonymous, or they may be linked to other information about the donor.

20.7 Medical Imaging:

- Medical imaging techniques are sometimes used to visualize parts of the human body.
- Examples include X-rays, CT scans, MRIs, and ultrasound.

20.8 Tests of Physical Fitness:

Kinesiology is the study of the mechanics, physiology, and psychology of body movement, function, and performance.

- Examples include test of strength, endurance, and flexibility.
- Safety must be the top priority.

Many different tests can be used to measure physical fitness levels:

- Cardiorespiratory fitness can be assessed using a 1-mile walking test, a 1.5- mile run test, or some other test of aerobic fitness.
- Measures of muscle strength and endurance include timed curl-ups, push-ups, pull-ups, flexed arm hangs, bench presses, leg presses, and grip tests (using a handgrip dynamometer).
- Flexibility can be measured using a sit-and-reach test (often measured with a flexometer) and other activities that stretch the lower back, hamstrings, or other muscle groups.
- Additional tests of fitness may assess agility, balance, coordination, speed, power, and reaction time.

20.9 Environmental Assessment:

- Environmental risk to human health can be assessed by trained observers with checklists and by laboratory tests.

The natural and built environments can have short- and long-term impacts on human health. Consider just a few of the many environmental factors that may affect the safety of a home:

- Is the entrance to the home accessible, or are there stairs or other barriers to access for people with mobility limitations? Is there adequate outdoor and indoor lighting?
- Are any stairs in the home loose or uneven? Do all stairs have handrails? Are all stairways free of clutter? Is any carpeting firmly affixed to the floor so that it will not slip?
- Does the bathtub or shower have a nonslip surface to prevent falls? Is the water heater set to prevent scalding and burns? Is the bathroom free of water damage, moisture, and mold?
- Do residents have reliable access to clean and safe drinking water?
- Is the kitchen free of pests and rubbish?
- Has the home been tested for toxic substances such as lead paint and asbestos? Is the home ventilated to prevent the buildup of radon gas? Are household chemicals, such as cleaning supplies, safely stored?
- Is the home equipped with working smoke alarms and carbon monoxide detectors?
- Does the home have adequate temperature control to prevent extreme heat and extreme cold?
- Are there sidewalks that facilitate safe walking near the home? Is the home located near a park, a playground, or another place where residents can safely engage in physical activity and recreation?

Similar lists of questions could be developed for schools, healthcare facilities, workplaces, and other locations

➔ This kind of studies require a lot of resources and that's why they are not commonly performed in our region

20.10 GIS (Geographic Information Systems):

A geographic information system (GIS): is a computer-based platform for mapping the locations of events, identifying spatial clusters, and testing complex spatial associations

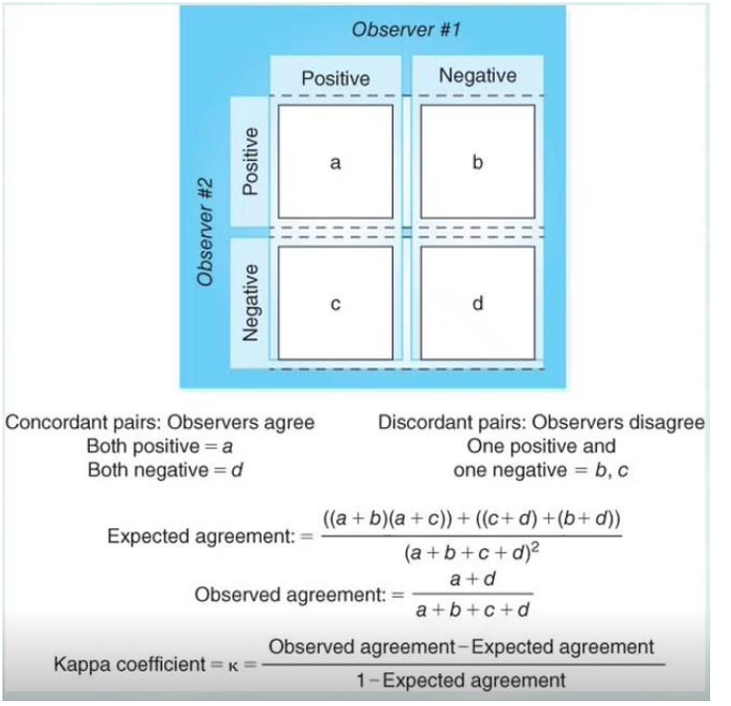
- Use a GPS (globe positioning system) to gather the latitude and longitude for key places, so that they can be mapped and incorporated into spatial analysis.

Example: To study the geographic distribution of coffee-houses that offer waterpipes (shisha) around universities in Jordan, this kind of studies offer possible interventions for policy makers to minimize the prevalence of water-pipe smoking among students.

20.11 Inter-Rater Reliability:

- Use the kappa statistic or another test to demonstrate that two assessors are making consistently valid measurements.

Statistical tests can be used **to determine the extent of agreement between two assessors who are evaluating the same study participants.** For example, a measurement known as **the kappa statistic** can indicate whether two radiologists examining the same set of X-rays reach the same conclusion about the presence or absence of a fracture more or less often than expected by chance. If the two radiologists agree as often as expected by chance, K = 0, If they agree on the interpretation of 100% of the X-rays shown to both of them, K = 1. If they agree more often than expected by chance, kappa will have a positive value somewhere between 0 and 1. Although complete agreement is rare, **a valid study will have a value of kappa that is close to 1**. Other measurements of inter-observer agreement or inter-rater agreement (also called concordance) can also be used to assess the validity and consistency of other assessment tools and procedures.



Chapter ~~21~~ 25: Secondary analyses

Some health research studies analyse existing clinical records, survey data, or population data rather than collecting new information.

21.1 Overview of secondary analysis:

- A **secondary analysis** is a study in which a researcher analyzes data collected by another entity.
- In a secondary analysis, **the researcher conducting the statistical analysis has not had (and does not have) any contact with the individuals whose data are being examined.**
  - ➔ Researchers should be aware and know how data was collected exactly.
  - ➔ Secondary analysis is often an excellent option for researchers with strong statistical skills but limited time and/or data collection resources.

21.2 Publicly Available Data Sets:

❖ **An anonymized data set (also called a deidentified data set):** is one that has been stripped of all potentially identifying information, such as names, street addresses, and personal identification numbers.

❖ **Deidentification:** is the process of removing potentially identifying information from a data file so that the data can be shared with others without violating the privacy of the individuals whose data are included in the file.

Many datasets are available online, such as a diversity of cross-sectional studies from the CDC:

- National Health and Nutrition Examination Survey (NHANES)
- National Health Interview Survey (NHIS)
- Behavioral Risk Factor Surveillance System (BRFSS)
  - ➔ The data may be downloadable to all, or there may be an application process for acquiring the relevant data.
  - ➔ Read all relevant supporting documentation and be aware of possible costs and authorship issues.

21.3 Private Data Sets:

- Professors (and others) may make **not-yet-analysed** data files available to students and other investigators.
- Ethical approval, careful review of the data collection methods, and frank conversations about authorship are required.
  - ➔ A request for access to a private data set is most likely to be granted when the new researcher has some existing connection to the original researcher.

21.4 Clinical Records:

- **Clinical records are a common source of data for case series.** Individuals working in clinical settings often can apply to gain access to patient records for research purposes.
- Ethical review of a research protocol is required before data are accessed (eg., **HIPAA** Act in USA).
  - ➔ **HIPAA**: a set of regulations about patient protection that must be carefully followed.
- Some data are electronic; others may be on paper and require data entry
- Records are often incomplete; absence of information in a file cannot be assumed to mean that the symptom or sign was not present in the patient.
- Some secondary analyses are conducted using deidentified records from registries. **A registry** is a centralized database containing information about people who have had a particular exposure or been diagnosed with a particular disease

- Most clinical sites require researchers to submit an application form to an oversight committee for review and approval prior to being authorized to access the data. The application must explain the goals of the study, the process that will be used to identify eligible patient records, the specific information that will be extracted from each patient's files, the steps that will be taken to protect the confidentiality of the data file, and the analysis plan. Applicants must also provide evidence of having successfully completed both research ethics training and specific instruction about patient privacy laws and policies. For example, researchers working with patient records in the United States must be prepared to comply with the **Health Insurance Portability and Accountability Act (HIPAA)** Privacy Rule.

21.5 Health Informatics, Big Data, and Data Mining:

- **Health informatics:** is the application of advanced techniques from information science and computer science to the compilation and analysis of health data.
- **Bioinformatics:** is the use of computer technologies to manage biological data.
  -> Bioinformatics typically focuses on analysis of molecular-level data (or, less often, tissue-level data). Clinical informatics and public health informatics usually focus on patient or population-level data. The tools of health informatics can be used to create novel data sets for research purposes.
- **Big data** refers to the analysis of data sets that are so large and complex that they require access to powerful hardware and special statistical software applications. These data sets may include data for many thousands or even millions of individuals from **data sources** such as:
  - Electronic health records (EHRs) or electronic medical records (EMRs), some of which use SNOMED CT (Systematized Nomenclature of Medicine Clinical Terms) as a standard terminology
  - Billing records, which often use ICD codes (Intonational Classification of Diseases codes) based on diagnoses or CPT codes (Current Procedural Terminology codes) based on procedures
  - Laboratory records, which often use LOINC codes (Logical Observation Identifiers Names and Codes)
  - Medication records, which often use NDC codes (National Drug Code identifiers)
  - Social media posts and other sources of information derived from the Internet
  - A diversity of other sources
- **Data mining** can be used to extract particular phrases from large sets of records

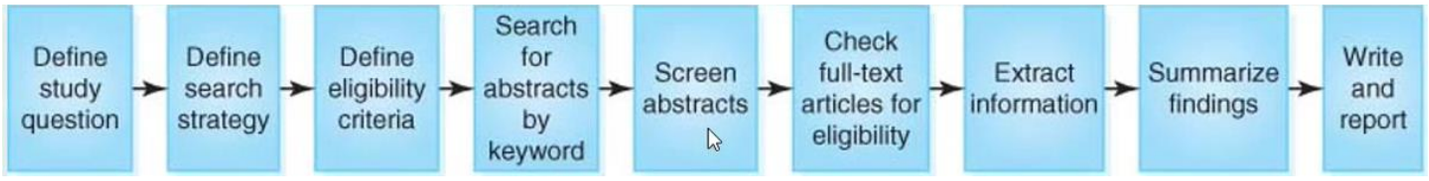
21.6 Ethics Committee Review:

- Any data file containing possibly identifiable information requires review by an ethics committee prior to beginning analysis.
- For all other data files, check with the appropriate committees about what review is required.
- Use of hospital records for research purposes always requires review by one or more research ethics committees. If the data for a secondary analysis come from a private source, then, prior to even looking at the data set, the analyst usually must obtain clearance from his or her own institution and perhaps also from the institution that houses the data. This permission must be secured prior to even looking at the data set. The application for permission to analyse existing data is often shorter than the application required for primary studies, and review usually can be expedited. It is better to err on the side of submitting an unnecessary proposal than to erroneously presume that a project is exempt from review without confirming the validity of this assumption.

- Most publicly available data, especially those collected by government agencies or federally sponsored researchers, were collected under protocols approved by one or several research ethics committees and then stripped of all personal identifiers prior to being shared. Additional approval by an ethics committee at the institution where the secondary analysis will be conducted is often not required when several conditions are met:
  - The data were collected after approval by a trusted organization's research ethics committee.
  - The data set contains no individually identifying information.
  - The data to be analysed are publicly available.
- However, researchers are responsible for becoming familiar with the requirements of their own institutions and ensuring that their work is compliant with all institutional policies. When there is any doubt about whether review is required, the institutional review board should be consulted.

The systematic review process:

1. Identification of an appropriately narrow study question.
2. Selection of a well-defined & valid search strategy [from where, & what keywords]
3. Screening of all potentially relevant articles [using the inclusion criteria].
4. Extraction of relevant information from all eligible articles.
5. Summarization of the findings of these articles



22.6 Search Strategy:

The first section to be written in the methodology of systematic review is search strategy

- The MeSH dictionary (available through PubMed.org) can help with focusing or expanding search terms by identifying the definitions of key terms as well as synonyms and related terms.

For example: a search for "health care costs" shows that synonyms for this term include "treatment cost" and "medical care costs." **Sub-headers** (sometimes called child terms) for "health care costs" include "direct service costs," "drug costs," "employer health costs," and "hospital costs." And "health care costs" fall under the headers (sometimes called parent terms of "health care economics and organizations," "health care quality, access, and evaluation," and "delivery of health care."

- Use **Boolean operators** such as AND, OR, and NOT.
- Confirm that the selected search string will capture several articles known to meet the eligibility criteria

➔ Examples of using Boolean operators:

- [a or b] → will find any abstract that includes "a" or "b" or both.
- [a AND b] → will yield only those abstracts that include both terms.
- [a NOT b] → will yield only those abstracts that include "a".
- More complex search strings may use parentheses, such as [a AND (b OR c)], which will find any abstract that includes both "a" and "b" or includes both "a" and "c."

FIGURE 22-3 Examples of Using Boolean Operators to Expand or Restrict the Number of Abstracts Identified in a Database

Search String	Approximate Number of "Hits" in PubMed
cancer	3.5 million
bladder cancer	70,000
schistosomiasis	25,000
"schistosomiasis"[Mesh]	21,000
Schistosomiasis mansoni	10,000
cancer AND schistosomiasis	1500
bladder cancer AND schistosomiasis	650
bladder cancer AND *Schistosomiasis mansoni*	40
bladder cancer OR schistosomiasis	90,000
bladder cancer NOT schistosomiasis	65,000
colorectal cancer	200,000
colorectal cancer AND schistosomiasis	150
bladder cancer AND colorectal cancer AND schistosomiasis	20
(bladder cancer OR colorectal cancer) AND schistosomiasis	800

Note: Because the PubMed database is constantly adding new abstracts, the numbers in this table will not exactly match the results of a new search.

❖ Understanding the language used by MEDLINE and other databases allows for the design of a database appropriate search string. **For example:**

- In MeSH language a "child" is defined as a person who is 6 to 12 years old. Individuals who are 2 to 5 years old are classified as "preschool children," and those who are 13 to 18 years old are "adolescents."
- A keyword search of [child]—that is, a search for the word "child" in all of the titles and abstracts of articles indexed in PubMed—will yield hundreds of thousands more hits than a search for ["child"[Mesh]] that only searches for articles indexed with "child" as a MeSH keyword. The researcher must be attuned to these particularities when designing search procedures.

To check the appropriateness of search terms, the researcher can identify several articles known to be relevant to the study question and then confirm that the search string captures all of those articles. If the search misses one or more of those key references, then the search strategy needs to be modified.

However, this process must not be used to exclude disliked articles, which would cause the inclusion bias that systematic reviews seek to minimize.

Once a validated system for identifying all of the potentially eligible articles is in place, the selected abstract databases are systematically searched for articles that might meet the inclusion criteria. If the topic is appropriately narrow, then keyword searches can often reduce the number of abstracts and/or articles that must be screened for eligibility to a reasonable number, often several hundred articles rather than many thousands of articles.

22.7 Search Limiters:

- ❖ Be careful about decisions to limit the search database screened, the languages or publication years of articles, and other choices that may reduce the number of articles identified by the search.
- ➔ For example, if you decided to search from midline only, or to choose articles in English language only, **you will have to justify** why you took these decisions because this can be considered as inclusion bias as they limit the number of identified articles.
- ➔ Or if you searched for broad keyword terms like “adult“ or “United States“ they may limit the search as many papers reporting about these populations may not include these terms.
- ➔ Researchers should be cautious about using the built-in filters available in some abstract databases. For example, PubMed allows researchers to use filters to restrict results to particular types of articles (such as clinical trials or reviews), particular species (such as human-only studies), and particular age groups (such as infants or adults aged 65+ years). These limiters work only if an article was indexed appropriately by the submitting journal. Because many articles about humans do not add “human” as a keyword and many studies do not include keywords for the ages of participants or the study design, the built-in limiters often exclude many studies that would otherwise be eligible for the review. It is usually better to use study-specific exclusion checklists to screen out abstracts that are ineligible rather than to artificially limit the number of screened abstracts using filters.

22.8 Eligibility criteria:

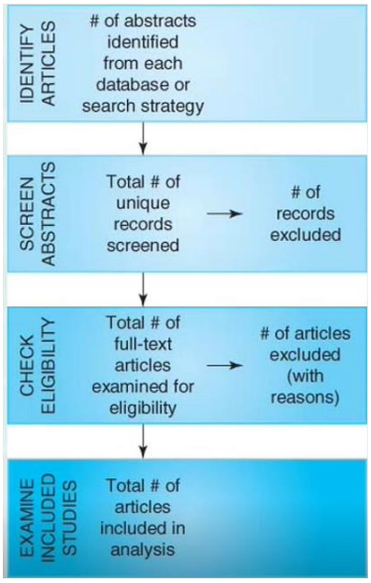
As the database searches are being conducted, each identified article’s title and abstract are reviewed to determine if they are eligible. The decision about an article’s eligibility for inclusion in a systematic review or meta-analysis is based on predetermined lists of inclusion and exclusion criteria. These eligibility criteria should ensure that all of the included studies pertain to the main research question.

- create both a list of inclusion criteria and a list of exclusion criteria prior to screening the articles to avoid bias.
- Be prepared to justify all criteria, especially those related to quality screening.
- Consider whether to use snowballing and the gray literature to expand the search.

snowballing: to look up every article cited in the eligible articles in order to identify other not indexed relevant articles.

The gray literature: articles which are available to reviewers not published yet.

hand searching: all articles in the tables of contents of selected volumes of relevant journals are scanned [any other source of information about your topic like magazines].



the figure summarizes the process of systematic review steps ->

- ❖ If the study question includes specific exposures, diseases or outcomes, and/or populations, the eligibility criteria should ensure that all of the included studies match those "EDPs."
- ❖ The eligibility criteria can also impose some requirements about the study design and sample size, if those restrictions are scientifically justifiable.
- ❖ Studies of causality are often appropriately limited to reviews of experimental studies, but reviews that are not focused on causation generally do not need to exclude observational studies.

22.9 Data extraction:

After identifying eligible articles, Data extraction tables list descriptive characteristics about the contents of these articles like: Author, location, year, population, sample size...etc.

- ➔ The **data extraction table allows for easy compilation & comparison of observations** relevant to the study question. And it's usually included in the search report.

22.10 Systematic review results:

- ❖ The researcher should record and report both statistically significant findings ($p < .05$) and statistically insignificant findings ($p \geq .05$) that are related to the main study question.

Why both?

- ➔ Let's say you studied 20 articles. when you say 10 articles has shown an association between specific variables, it gives an idea that there is an association, but when you mention both; that 10 articles did show an association and other 10 didn't show any association, this would be more clear that it's 50% of all articles did show an association.
- ➔ **So the studies that do not show statistically significant results <u>are just as valuable as those that find a significant association</u>.** Actually, one of the primary contributors of the systematic reviews to the health literature is the ability to identify both: areas of consensus and areas of (disagreement & uncertainty).

- ❖ **Publication bias** occurs when articles with statistically significant results are more likely to be published than those with null results.

(Proving that publication bias has occurred may not be possible, but the presence of consensus should be cautiously interpreted when only limited number of studies are published on the topic or the results are mixed.)

We'll continue this chapter in the next lecutre...

22.11 Meta-Analysis

Meta-analysis (tertiary analysis) **combines into one summary statistic the results of several high-quality quantitative studies that used similar methods to collect and analyze their data.** The inclusion criteria for meta-analyses are usually more restrictive than they are for general systematic reviews; so systemic review and meta-analysis are separated procedures!

The meta-analysis process:

- **Use a systemic search strategy to identify relevant articles.**
- **Carefully read each study.**
- **Assess the quality and comparability of each study.**
- **Extract statistical results from each of the eligible studies.**
- **Combine comparable statistical results into one summary statistic.**

22.12 Pooled Analysis

Homogeneous (similar) studies can be combined into a summary statistic, but caution should be used if the studies are heterogeneous (dissimilar).

The amount of variability in the measure between studies can be examined using a Cochran's Q statistic for homogeneity and the I^2 statistic.

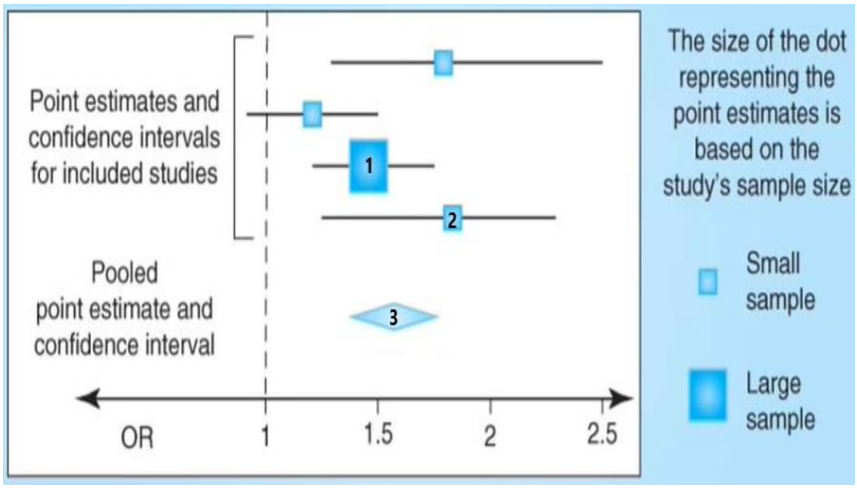
- There are two main choices of models to use for meta-analysis:

1. **A <u>fixed effects model</u> can be used to create a pooled estimate when the studies are fairly homogeneous.** (There is little variability)
2. **A <u>random effects model</u> is required when the tests of heterogeneity show that the included studies are dissimilar.** (There is considerable variability)

** The results from studies using different study designs, different interventions, or dissimilar population groups should not be pooled

22.13 Forest plots and funnel plots

A forest plot displays the contributing studies and the summary measure for a meta-analysis. >>>>>>>>>

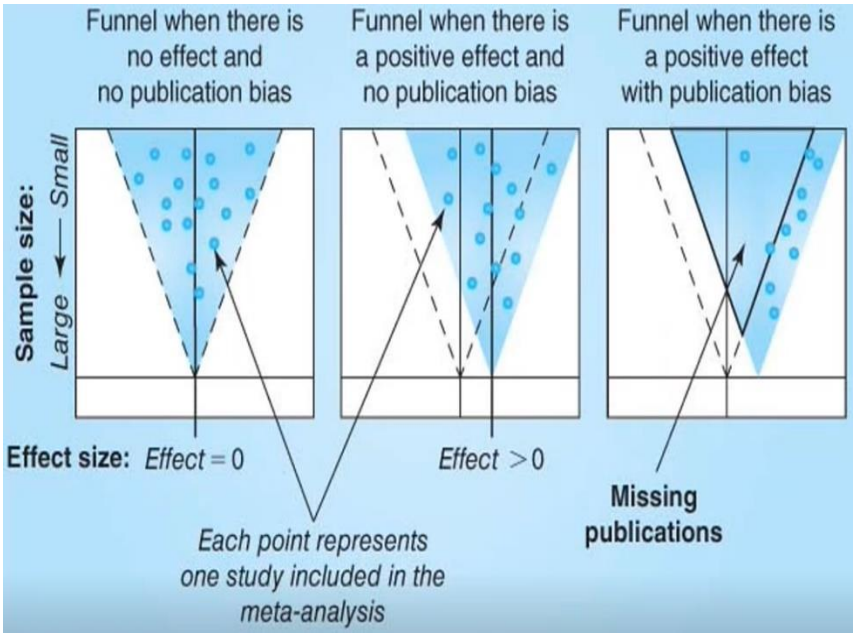


Effect size is the magnitude of the difference in the value of a statistic in independent populations. Many types of statistics can quantify effect sizes, including **odds ratios (OR),** difference in means measures…

Each square represents a different study whose sample size is represented by the size of the square; [2] is smaller but more risk than [1].

[3] A diamond shape represents what the potential risk collectively based on all those studies (the sum of their results).

A funnel plot visually displays the likelihood of studies missing from the analysis because of publication bias. >>>>>>>>>



If no publication bias has occurred, the points for the included studies will form a triangle. If publication bias has reduced the number of publications with statistically insignificant results, part of the triangle will be missing. In that situation, the pooled estimate is likely to have overestimated the true effect size.