SCIENTIFIC MEDICAL RESEARCH

WRITER: غادة الزعبي ، رؤى هديب

LECTURE: 13 + 14

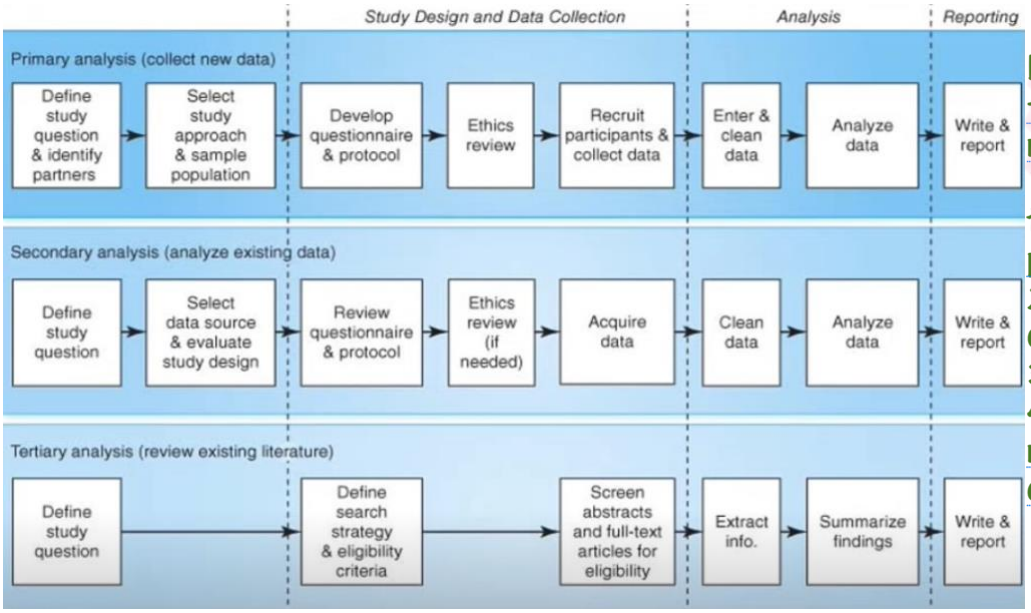
Chapter 15 : Research Protocols

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

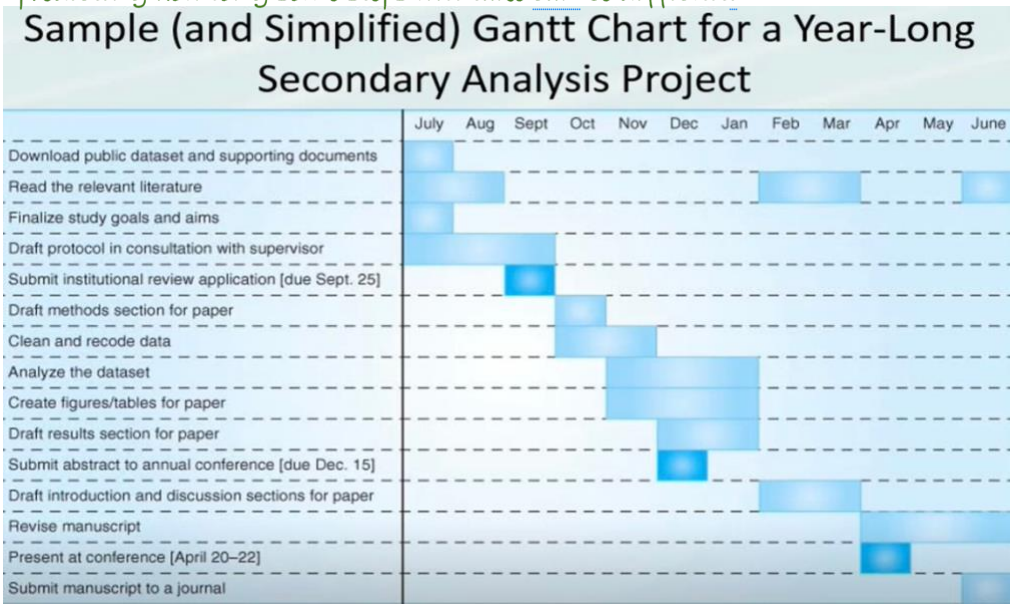
15.2 Research Timelines

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.

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

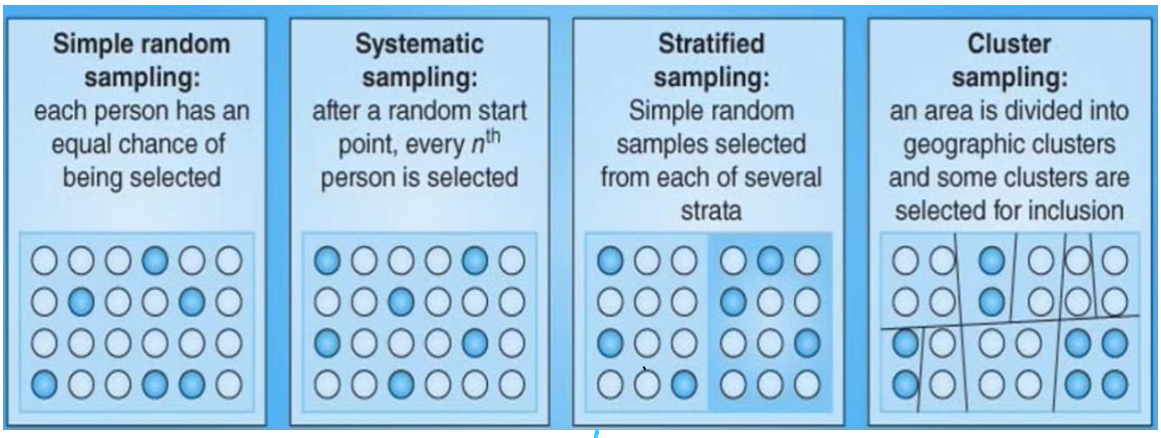
→ Source of population (sampling frame): a list of particular people from whom a sample population can be drawn.

→ 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.

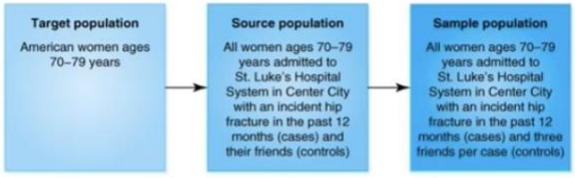
16.6 Populations for Case-Control Studies

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



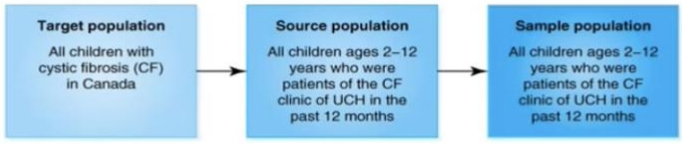
Study approach	Case-control study
Study question	What are the risk factors for hip fractures in adult women in the United States?
Data collection method	Participants will be interviewed in person or by telephone.
Target population	Women ages 70–79 living in the United States
Source population	All women 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.

Chapter 17 : Sample Size Estimation

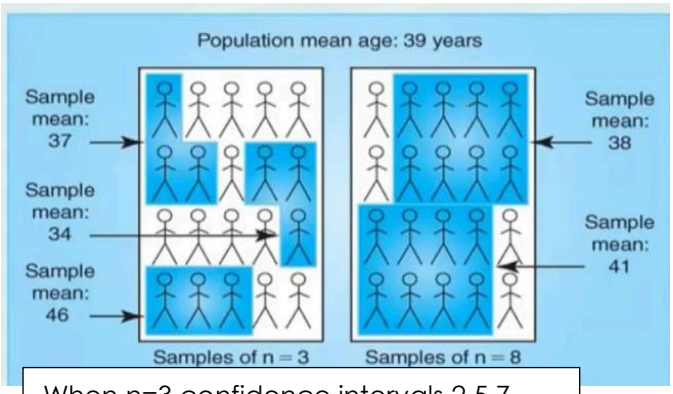
17.1 Importance of sample Size

- 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

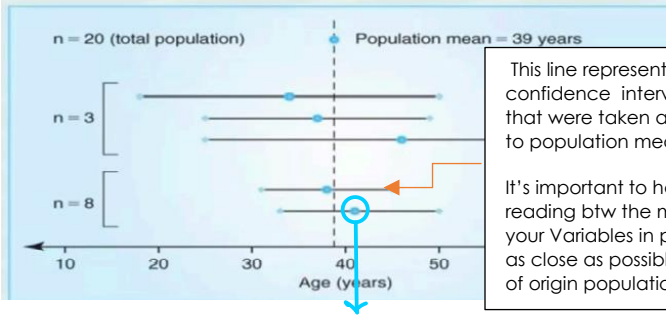
- Larger sample size produces narrower confidence intervals (an estimate of how close to the population value a sample of a particular size is expected to be) for statistical measures.
- Larger sample sizes yield more statistically significant results.



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

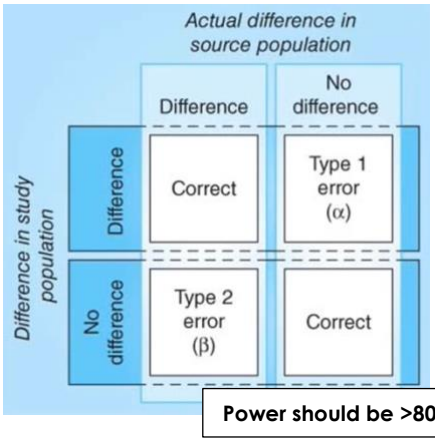
FIGURE 17-4 Sample Size Estimates for a Case-Control Study

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.