



WHAT IS RESEARCH?



A Process of Systematic, Scientific Data

- Collection
- Analysis &
- Interpretation

RESEARCH is

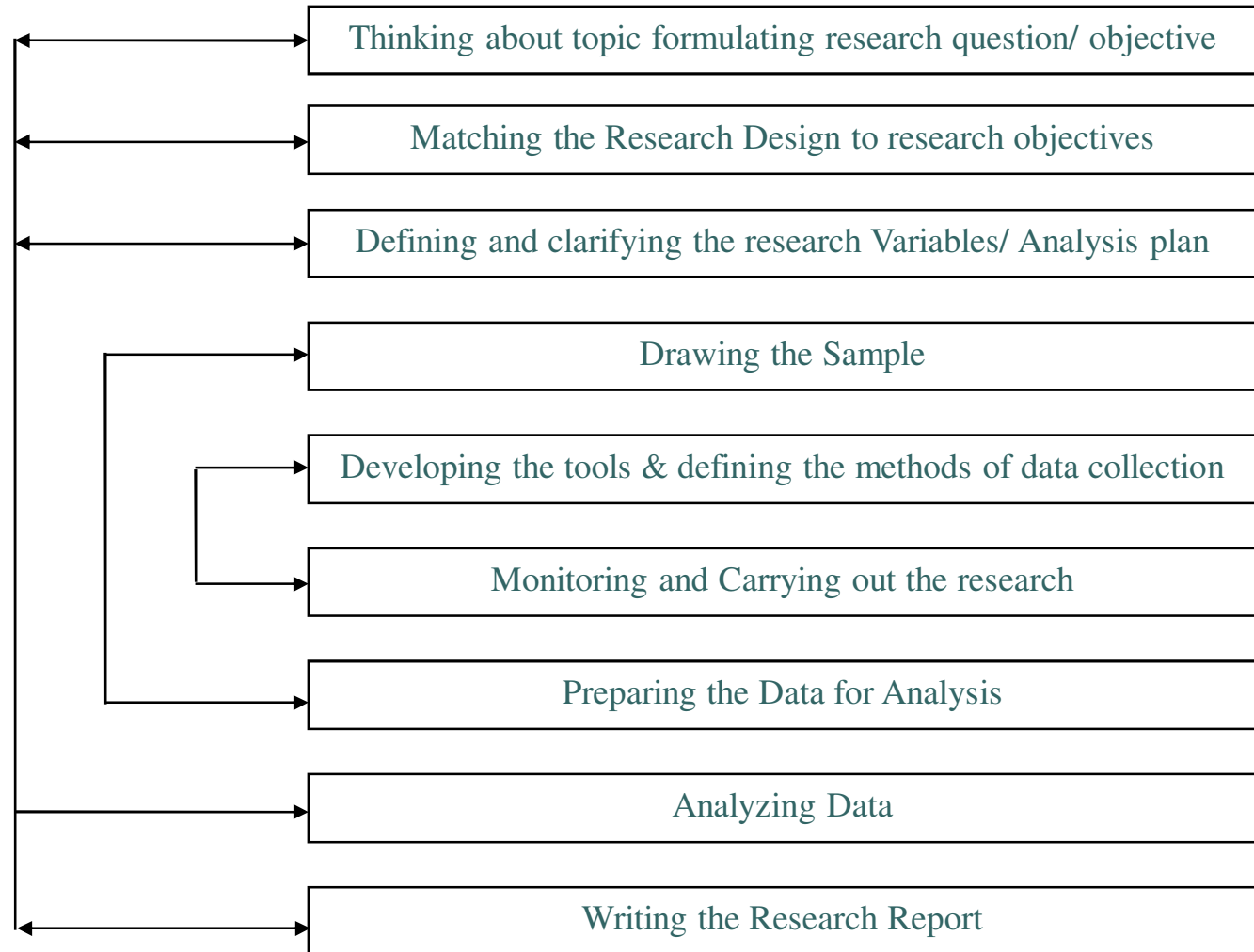
**So as to find Solutions to
a problem.**



Types of Research

- **Qualitative**
- **Quantitative**

STEPS IN DESIGNING AND CONDUCTING RESEARCH





SELECTING A RESEARCH TOPIC



CRITERIA FOR SELECTING A RESEARCH TOPIC

- *Relevance*
- *Innovation*
- *Feasibility*
- *Acceptability*
- *Cost-effectiveness*
- *Ethical consideration*

Priority Ratings for Research Proposals



How big is the problem?
 How important is it to look for relevant solutions to it ?
 Are solutions to it available, if so how effective have they been proved?
 Is the problem of import in our local set up

CRITERIA FOR SELECTION							
Proposed Topic	Relevance	Innovation	Feasibility	Acceptability	Cost Effectiveness	Ethical	TOTAL SCORE

Priority Ratings for Research Proposals



Proposed Topic	CRITERIA FOR SELECTION						
	<i>Relevance</i>	<i>Innovation</i>	<i>Feasibility</i>	<i>Acceptability</i>	<i>Cost Effectiveness</i>	<i>Ethical</i>	<i>TOTAL SCORE</i>

Incase the topic has been researched what new are you looking at

Priority Ratings for Research Proposals



Proposed Topic	CRITERIA FOR SELECTION						TOTAL SCORE
	Relevance	Innovation	Feasibility	Acceptability	Cost Effectiveness	Ethical	

Consider the resources that are required to carry out the project.

Priority Ratings for Research Proposals



Proposed Topic	CRITERIA FOR SELECTION						
	Relevance	Innovation	Feasibility	Acceptability	Cost Effectiveness	Ethical	TOTAL SCORE

research a topic which
has the interest and
support of the authorities

Priority Ratings for Research Proposals



Proposed Topic	CRITERIA FOR SELECTION						
	<i>Relevance</i>	<i>Innovation</i>	<i>Feasibility</i>	<i>Acceptability</i>	<i>Cost Effectiveness</i>	<i>Ethical</i>	<i>TOTAL SCORE</i>

Whether the resources of time, money and manpower being invested in the study are worthwhile

Priority Ratings for Research Proposals



Proposed Topic	CRITERIA FOR SELECTION						
	<i>Relevance</i>	<i>Innovation</i>	<i>Feasibility</i>	<i>Acceptability</i>	<i>Cost Effectiveness</i>	<i>Ethical</i>	<i>TOTAL SCORE</i>

- Cultural sensitivity must be given careful consideration.
- Informed consent ensured.
- Will treatment be given to individuals identified during study who require treatment?

Priority Ratings for Research Proposals



Proposed Topic	CRITERIA FOR SELECTION						
	<i>Relevance</i>	<i>Innovation</i>	<i>Feasibility</i>	<i>Acceptability</i>	<i>Cost Effectiveness</i>	<i>Ethical</i>	<i>TOTAL SCORE</i>

Rating Scale:
 1 = low,
 2 = medium,
 3 = high
Total Score out of 18



Literature Search



Literature Search:

What?

- Allows one to search in a purposeful and systematic manner, through a range of literature or information relevant to ones particular field, and to hone in on material relevant to ones interest and objectives.



A Literature Search

Why?

1. To keep up with the latest developments in your field.
2. To learn more about some topic.
3. To document important facts and ideas you wish to research in light of previous work done on it.
4. To understand your data in the context of what is already known.
5. To provide your readers with sources they can consult on their own.



Searching Sources

INFORMATION SOURCES	DATABASE SEARCH	HAND SEARCH	
		INDEXES	INFORMATION SOURCES
Journal article	MEDLINE & CANCERLIT or EMBASE (Internet)	Index Medicus Excerpta Medica	Key journal issue Reference lists of articles
Current journal articles	Current Contents or PREMEDLINE PUBMED	Current Contents/Clinical Medicine	Current journal articles
Research projects	PDQ (Physician Data Query file-ongoing trials)	Pakistan Government Research Documents	PMRC directory of research
Conference, congress, & meeting proceedings	Conference papers Index Directory of published proceedings (CPSP)	Conference Papers Index Directory of published proceedings	Program/proceedings of various societies (PPA, etc)
Researchers & research organizations	Research Centers and Services Directories	Research Centers Directory	First authors of relevant articles



Means of Literature Search

INTERNET

- Access to a massive pool of information related to biomedical and clinical source.
- Can link with library catalogues, online databases, like MEDLINE and direct access to ever increasing number of biomedical journals.
- Retrieval of data from a range of organizations universities, research establishments and hospitals.

Many services are available free of charge.



Search Strategy on Internet

- ❖ **Summarize your topic in one or two sentences.**
- ❖ **Identify the unique ideas or concept associated with your topic.**
- ❖ **Choose appropriate keywords for each concept.**
- ❖ **Establish the relationship between each keyword and concept.**



Example Pub Med Sources

- PubMed is derived from two words, Publications, and Medical.
- It is a project of the National Institute of Health, National Library of Medicine.
- For more details of PubMed, you may visit.

<http://www.ncbi.nlm.nih.gov/PubMed/>

This is how the PubMed Window would look like!



Entrez-PubMed - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss

Address <D:\CPSP Research Files\PubMed\Entrez-PubMed.htm> Go Links

NCBI National Library of Medicine PubMed

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM

Search PubMed for Go Clear

Limits Preview/Index History Clipboard

About Entrez

Entrez PubMed
Overview
Help | FAQ
New/Noteworthy

PubMed Services
Journal Browser
MeSH Browser
Single Citation
Matcher
Batch Citation Matcher
Clinical Queries
Cubby

Related Resources
Order Documents
Grateful Med
Consumer Health
Clinical Alerts
ClinicalTrials.gov

Privacy Policy

- Enter one or more search terms, or click [Preview/Index](#) for advanced searching.
- Enter [author names](#) as smith jc. Initials are optional.
- Enter [journal titles](#) in full or as MEDLINE abbreviations.

PubMed is the National Library of Medicine's search service that provides access to over 11 million citations in MEDLINE, PreMEDLINE, and other related databases, with links to participating online journals.

Explore beyond PubMed

Coffee Break Can't tell BLAST from VAST? [Coffee Break](#) demonstrates online bioinformatic tools available at the NCBI through live searches and interactive tutorials.

New Cubby feature! The Cubby provides you with a Stored Search feature to store and update searches. It also allows you to customize your LinkOut display to include or exclude links to providers. See [Help](#) and [FAQ](#) for additional information.

[Write to the Help Desk](#)

Internet

Start Microsoft PowerPoint - [Pu...] Entrez-PubMed - Mic...

7:50 PM



Assume that I am trying to work a research question on incidence of hypertension in diabetic patients with nephropathic complications.

When I enter a common disease name, like diabetes, I get as many as these search results! More than 162 thousand.... just unmanageable.

So the best idea would be to go for a more “restricted search”.....

Entrez-PubMed - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites

Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=PubMed>

PubMed Search

NCBI National Library of Medicine

PubMed Search PubMed

For diabetes

Go

Display Summary Save Text Order Details

Show: 20 Items 1-20 of 162306

1: [Bhatnagar D, Morgan J, Siddiq S, Mackness MI, Miller JP, Durrington P](#)
Outcome of case finding among relatives of patients with know



Say, “diabetes” and “hypertension”. Observe that the number of search results has fallen rather steeply. Less than a tenth!

I see only about 16000 results! That is much better, yet, who would read all of these articles?

The screenshot shows the NCBI PubMed website interface. The search bar contains the text "diabetes hypertension", which is circled in red. Below the search bar, the number of results is displayed as "16057", also circled in red. The interface includes a navigation bar with tabs for PubMed, Nucleotide, Protein, Genome, Structure, and PopSet. A sidebar on the left lists various services like Journal Browser, MeSH Browser, and Single Citation Matcher. The main content area displays a list of search results, with the first two entries visible:

- ☐ 1: [Bhatnagar D, Morgan J, Siddiq S, Mackness MI, Miller JP, Durrington PN.](#)
Outcome of case finding among relatives of patients with known heterozygous familial hyp
BMJ. 2000 Dec 16;321(7275):1497.
[Record as supplied by publisher]
PMID: 11118175
- ☐ 2: [Zabetakis PM, Nissenson AR.](#)
Complications of Chronic Renal Insufficiency: Beyond Cardiovascular Disease.
Am J Kidney Dis. 2000 Dec;36(6 Suppl 3):S31-S38.

So I decide to add “nephropathy” to that. Now the search results are a much better off 3010.

I may similarly add more search words to improve the number of results.

NCBI

PubMed Nucleotide Protein Genome Structure PopSet

Search PubMed for diabetes hypertension **nephropathy** [Go] [Clear]

Limits Preview/Index History Clipboard

Display Summary [v] Save Text Order Details Add to Clipboard

Show: 20 Items 1-2 **of 3010** Page 1 of 151

☐ 1. [Zabetakis PM, Nissenson AR.](#)
Complications of Chronic Renal Insufficiency: Beyond Cardiovascular Disease.
Am J Kidney Dis. 2000 Dec;36(6 Suppl 3):S31-S38.
[Record as supplied by publisher]
PMID: 11118136

☐ 2. [Hong CY, Chia KS, Ling SL.](#)
Urinary protein excretion in Type 2 diabetes with complications.
J Diabetes Complications. 2000 Sep 1;14(5):259-265.

About Entrez

Entrez PubMed
Overview
Help | FAQ
New/Noteworthy

PubMed Services
Journal Browser
MeSH Browser
Single Citation
Matcher
Batch Citation Matcher
Clinical Queries



Search strategy thru articles

Journal articles

Both primary and review publications end with a section called “Literature Cited” or References”, which provides the complete bibliographic information on every source used, these can be used for further literature search.



Research Objective



Research Objectives

- An objective is an intent of what the researcher wants to do stated in clear measurable terms.”



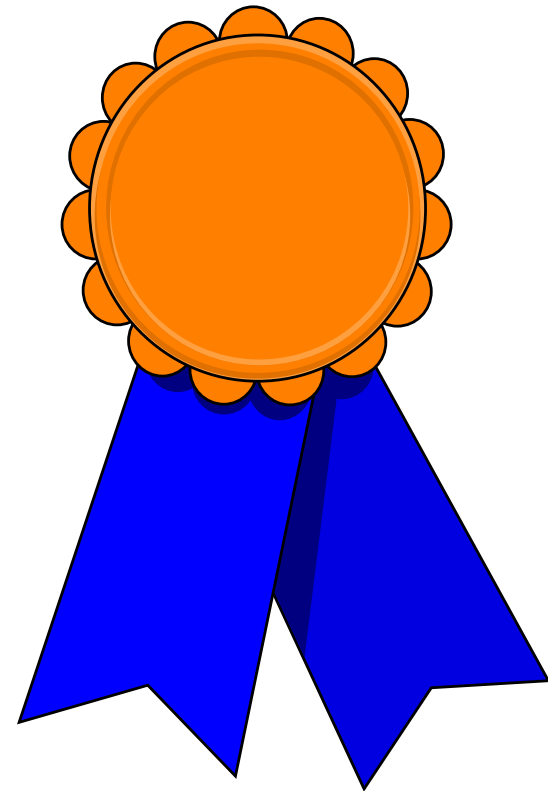
Importance of Research Objectives

1. Brings focus to the study.
2. Avoids collection of unnecessary data.
3. Determines an appropriate study design.
4. Helps determine analysis plan.



A Good Objective ensures that:

What is to be measured is clearly stated, be it a measure of frequency, or Association in the population of interest.

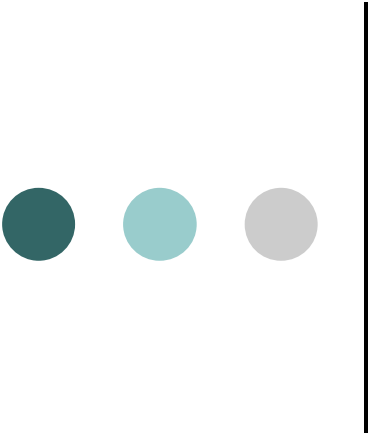




Examples

Objectives:

- 1) To determine the frequency of anemia in pregnant women visiting Tertiary care facilities of Sindh.
- 2) To determine association between maternal smoking and LBW.
- 3) To compare the effectiveness of dressing A vs. dressing B in patients presenting with infected wounds of the foot.

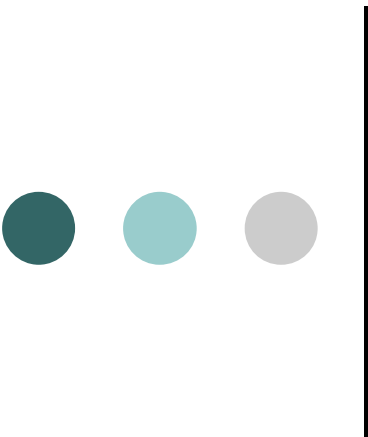
A decorative graphic consisting of three colored circles (dark teal, light teal, and grey) arranged horizontally, followed by a vertical black line.

Operational Definition



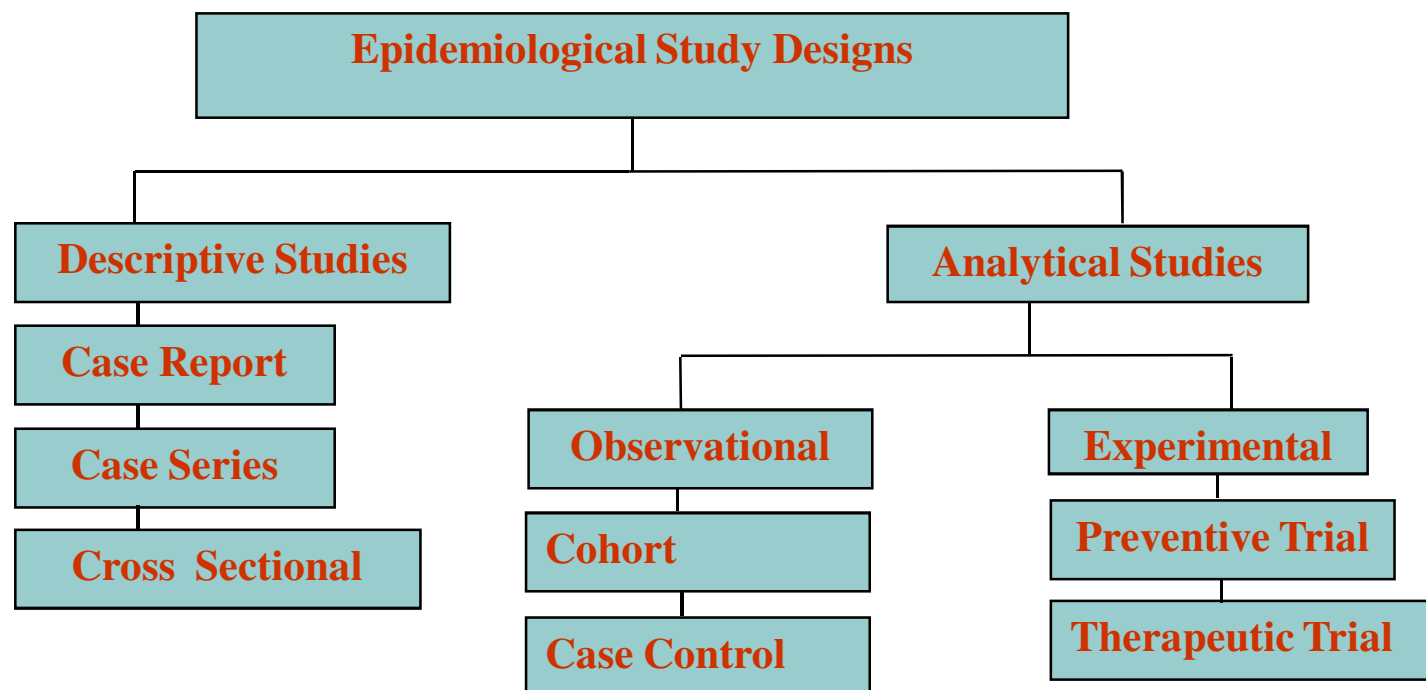
Operational Definitions

- Is the definition of the exposure and outcome variables of interest in context to objective in a particular study and their means of measurement/determination.
- Examples:
 - Anemia
 - Effectiveness
 - PPH
 - Wound healing



Study Designs







DESCRIPTIVE STUDIES

Descriptive studies involve the systematic collection and presentation of data to give a clear picture of a particular situation and can be carried out on a small or large scale.

- **Case Report**
- **Case series**
- **Cross Sectional Survey**



COMPARATIVE or ANALYTICAL STUDIES

- An ANALYTICAL STUDY attempts to establish association or determine risk factors for certain problems. This is done by comparing two or more groups, with or without the outcome of interest/exposure of interest.

Types

- Observational
- Experimental



Case Report

- A detailed report by a physician of an unusual disease in a single person.
- Classical example is that of a single case reported in Germany in late 1959 of a congenital malformation affecting the limbs and digits.
- More cases were reported in the following years. In 1961 a hypothesis was put forward that thalidomide, a sleeping pill, was responsible for congenital malformations.
- Subsequent analytic studies confirmed the link between the drug and congenital malformation.



Case Series

- When several unusual cases all with similar conditions are described in a published report, this is called a Case Series.
- A case series does not include a control group.

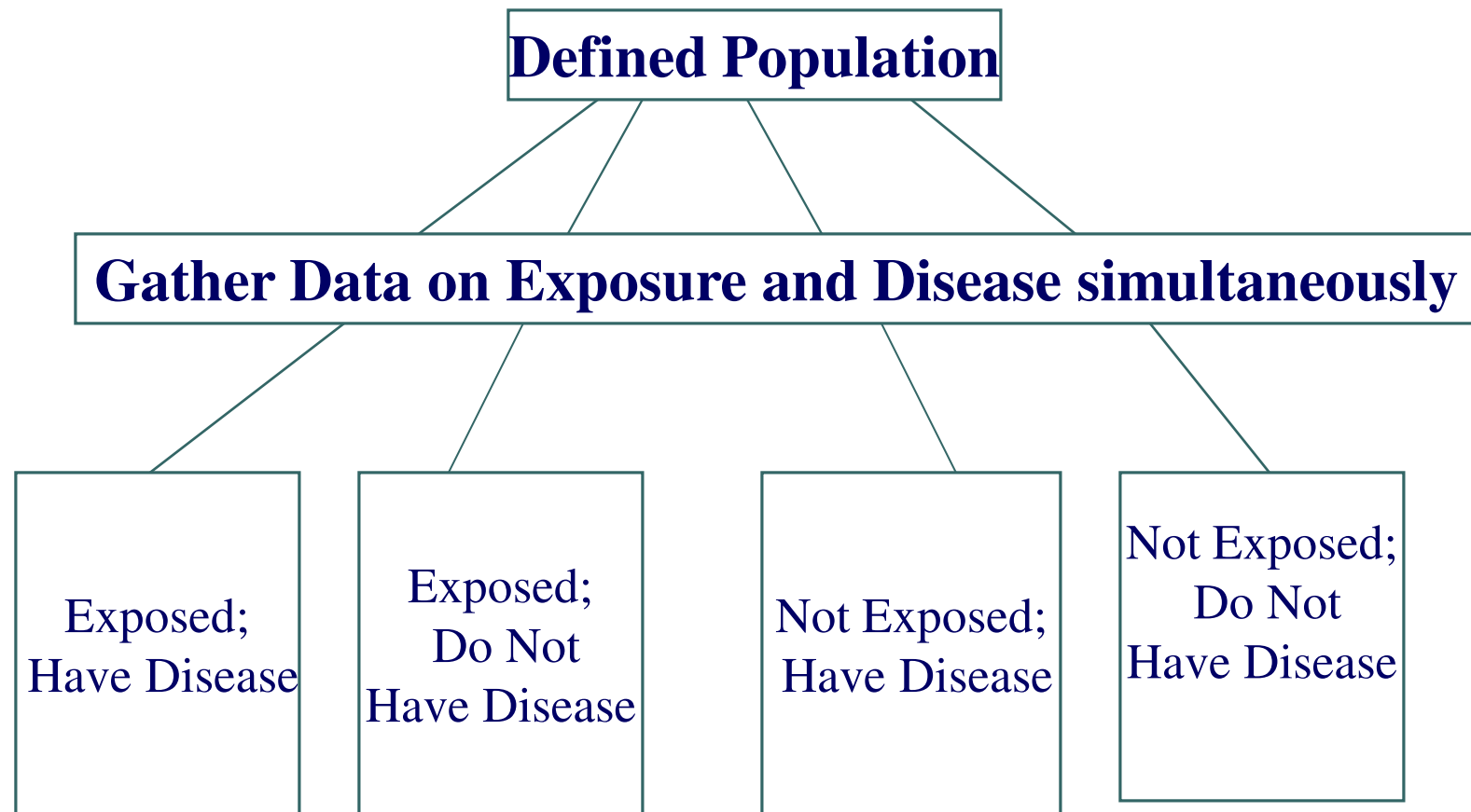


Cross-Sectional Studies

- A cross-sectional study is a survey of a defined population at a single point in time.
- Also called “Prevalence Studies”.
- Determines the “burden of disease” in a population.
- Exposure & outcome are determined simultaneously.



Design of a Cross-Sectional Study





Cross-Sectional Surveys-Advantages



- Fairly quick and easy to perform.
- Inexpensive.
- Useful for determining the *prevalence* of disease for a defined population and can also measure factors leading to it subsequent to group formations.



Cross-Sectional Surveys - Disadvantages

- *Provide only a “snapshot” in time.*
- The data about both the exposure to risk factors and the presence or absence of disease are collected simultaneously, hence it is *difficult to determine temporal relationship of a presumed cause and effect.*



Cohort Studies

- A cohort is a group of people who have something in common (a characteristic or characteristics suspected of being a precursor to or risk factor for a disease) and who remain part of a group over a period of time.



Types of Cohort Studies

- Prospective Cohort Studies
- Retrospective Cohort Studies



Prospective Cohort Studies

The investigator assembles the study groups in *the present time*, collects baseline data on them and then continues to collect data for a period that can last many hours to years.

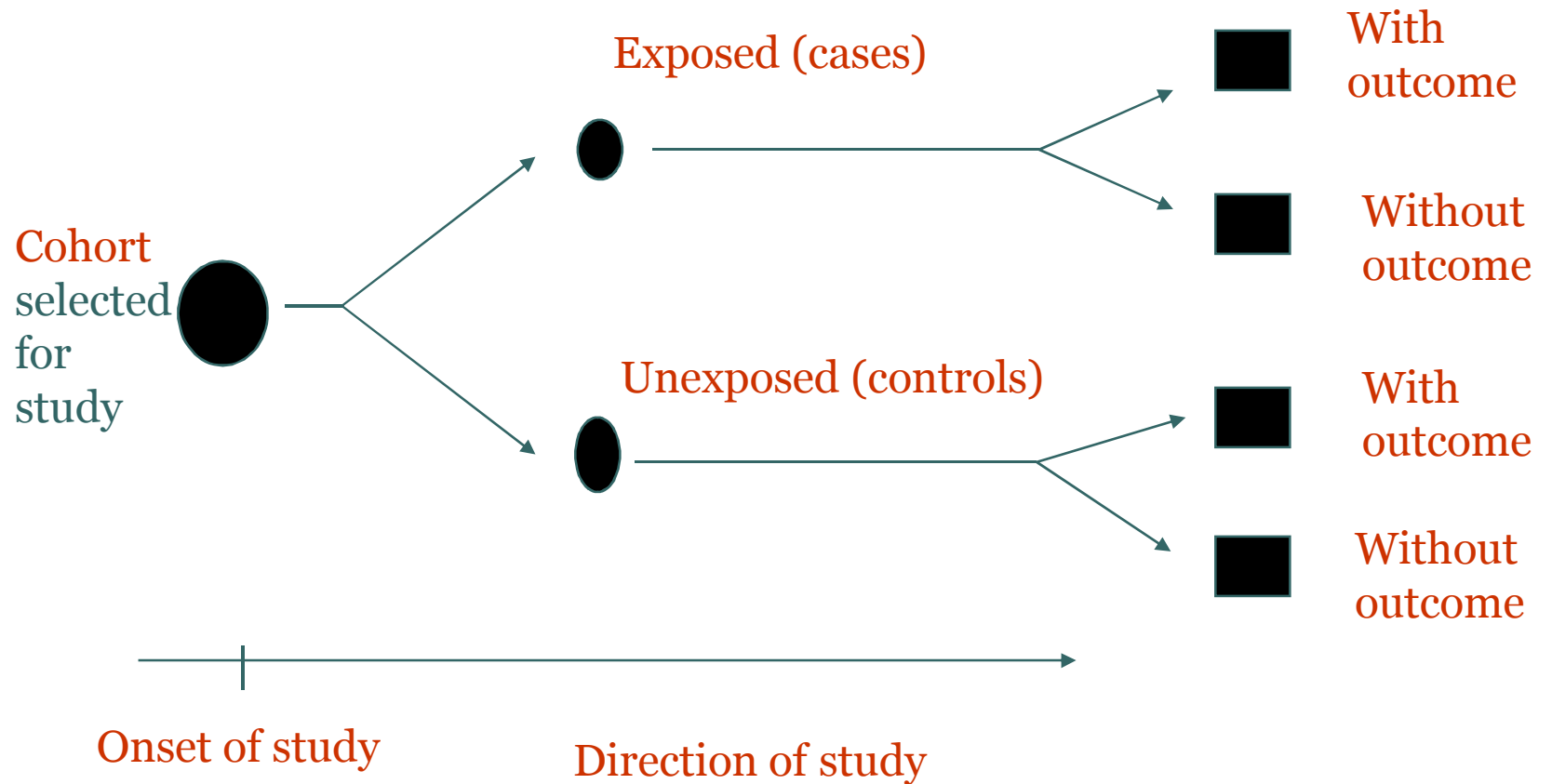


Retrospective Cohort Studies

- The investigator goes back into history to define a risk group (e.g. *those children exposed to x-rays in utero vs. those not), and follows the group members up to the present to see what outcome (cancer) have occurred

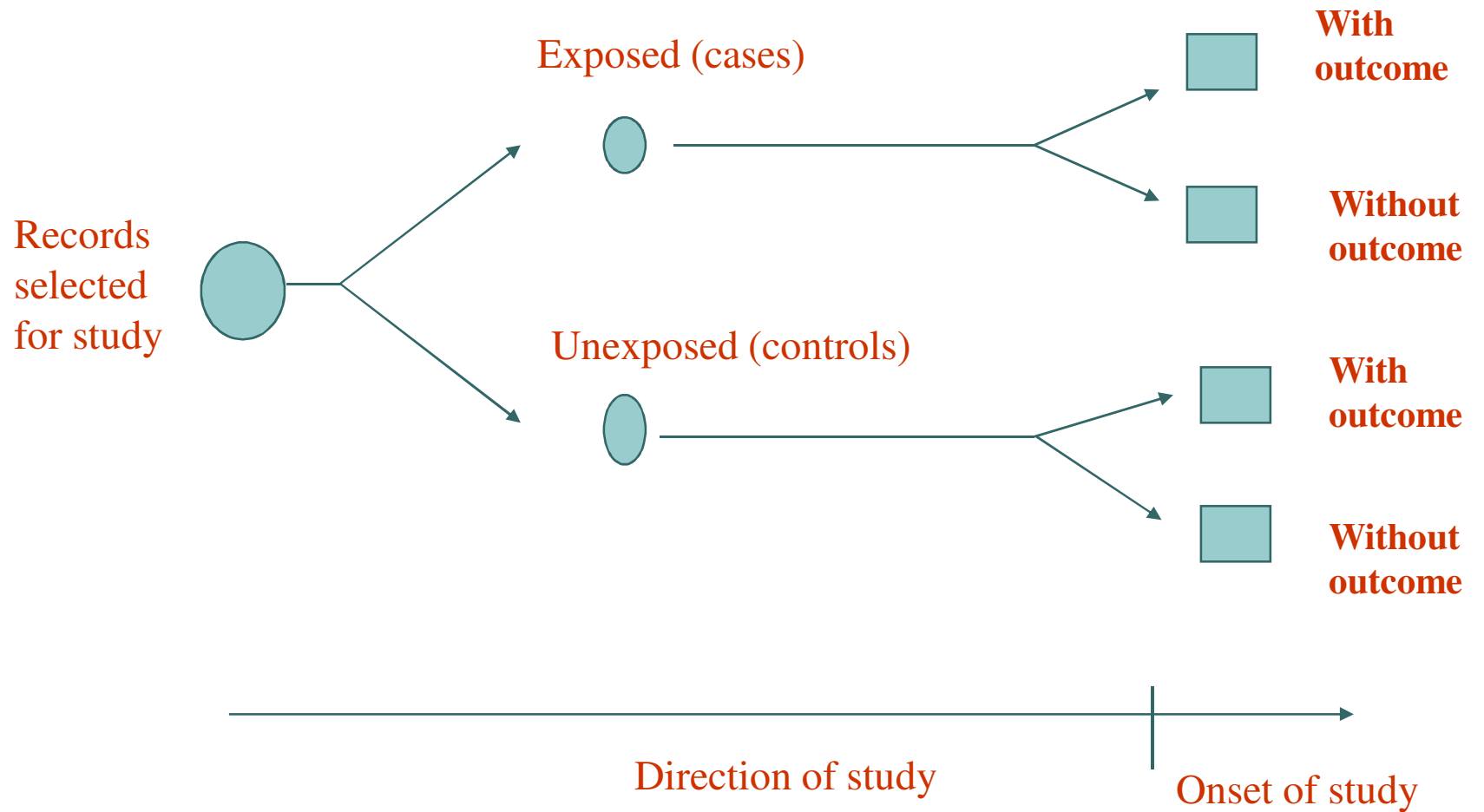


Cohort Study – Prospective





Retrospective Cohort Studies





Prospective Cohort Studies - Advantages

- Because they are longitudinal, are the study of choice for
 - ❖ Establishing causes of a condition (temporal relation)
 - ❖ Allows for measurement of incidence
 - ❖ Study of multiple effects of a single exposure



Prospective Cohort Studies- Disadvantages

- *With diseases that develop over a long period of time, or with conditions that occur as a result of long-standing exposure, many years are needed and hence:
 - ❖ High costs
 - ❖ Long wait until results are obtained
 - ❖ Loss to follow
- Are problematic when disease or outcome is rare
For example, studying the risk factors/ clinical features associated with carcinoid tumors (very slowly growing tumors)



Retrospective Cohort Studies

Advantages:

- Less expensive
- Completed in much shorter time than a prospective study

Disadvantages:

- The quality of data collection is not as good As records generated for clinical purposes and not for research
- Because of a large number of biases associated with these studies, carry less weight in establishing a cause than prospective studies.

Classical Example of a Cohort Study – The Framingham Study



- Framingham is a town in Massachusetts, about 20 miles from Boston
- Residents were considered eligible if they were between 30-62 years of age
- A sample size of about 5000 men and women
- Many exposures were defined such as smoking, obesity, elevated blood pressure, elevated cholesterol levels, low level of physical activity etc.
- Outcome: New coronary events – identified by examining the population every two years and by daily surveillance of hospitalization.

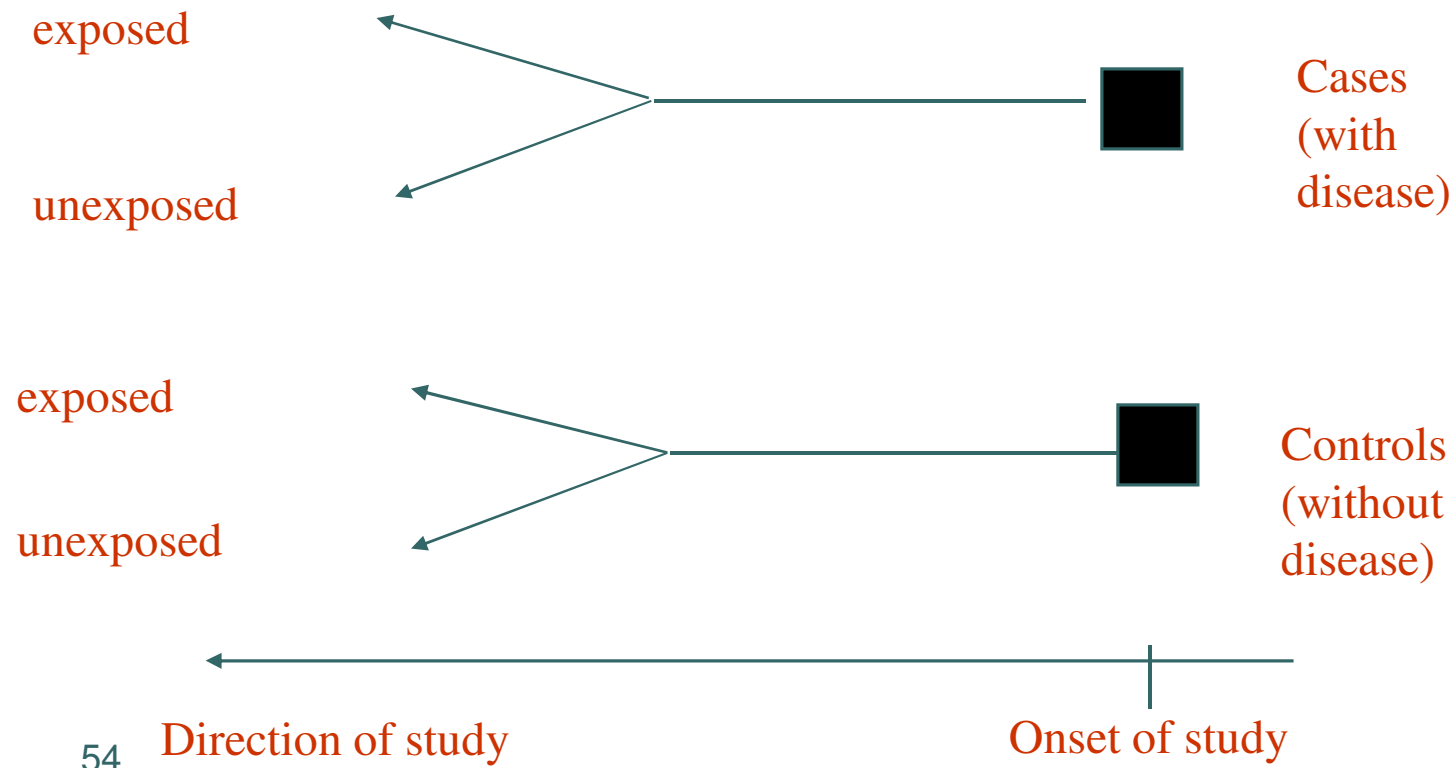


Case Control Study

- The investigator selects the case group and the control group on the basis of the outcome (i.e. having the disease of interest vs. not having the disease of interest)
- Cases and controls are assembled and are questioned or their medical records are consulted regarding past exposure to risk factors



Case Control Studies





Case Control Studies

Advantages

- Inexpensive
- Quick
- Especially useful when the disease being studied is rare or for a condition which develops over a long time
- Can evaluate multiple etiologies for one outcome

Disadvantages

- Recall bias
- Selection Bias



Case-Control Studies - Selection of Cases

Cases can be selected from a variety of sources:

- Hospital patients
- Patients in Physician's practices
- Registries



Selection of Controls

Controls may be selected from

- Non-hospitalized persons living in the community similar to cases
- Hospitalized patients admitted for diseases other than that for which cases are admitted



Problems with Case Control Study

Selection Bias

- In 1929, Raymond Pearl at John Hopkins, Baltimore conducted a study to test the hypothesis that tuberculosis protected against cancer
- He selected 816 cases of cancer from 7500 consecutive autopsies
- He also selected 816 controls from others on whom autopsies had been carried out at John Hopkins
- Of the 816 **CASES** (with cancer), 6.6% had TB
- Of the 816 **CONTROLS** (without cancer), 16.3% had TB
- From the finding that the prevalence of TB was considerably higher in the control group, *Pearl concluded that TB was protective against cancer*
- *Was Pearl's conclusion justified?*



Problems with Case Control Studies

“Pearl’s Study”

- No!! At the time of the study, TB was one of the major reasons for hospitalization at Johns Hopkins Hospital
- Pearl thought that the control group’s rate of TB would represent the level of TB in the general population; but because of the way he selected the controls, they came from a pool that was heavily weighted with TB
- *The way the controls are selected is a major determinant of whether a conclusion is valid or not*



Problems with Case Control Studies

Coffee-drinking and Cancer of the Pancreas in Women*

- Cases were white cancer patients from 11 Boston and Rhode-Island hospitals
- Controls were patients from GI Clinics
- McMohan found that coffee consumption was greater in cases than controls
- Controls were patients who had reduced their coffee consumption because of Physician's advice
- The controls level of coffee consumption was not representative of the general population
- When a difference in exposure is observed between cases and controls we must ask "*Is the level of exposure observed in the controls really the expected level in the general population.*"



Recall Bias

Individuals who have experienced a disease or other adverse health events tend to think about possible causes & thus are likely to recall histories of exposure differently as compared to controls.



INTERVENTIONAL / EXPERIMENTAL STUDIES

- The researcher manipulates a situation and measures the effects of the manipulation amongst two groups, one in which the intervention takes place (e.g. treatment with a certain drug) and another group that remains "untouched" (e.g., treatment with a placebo) .



EXPERIMENTAL STUDY

- Only type of study design that can actually prove causation
- Individuals are randomly allocated to at least two groups. One group is subjected to an intervention, while the other group(s) is not.
- The outcome of the intervention (effect of the intervention on the dependent variable) is obtained by comparing the two groups.



Characteristic features of Experimental Study



- Assignment of exposure (intervention) by the researcher
- An intervention and a comparative group
- Random allocation



SAMPLING



SAMPLING

A **sample** is a sub set of the **population**, with all its inherent qualities. Inferences about the population can be made from the measurements taken from a sample, if the sample is truly representative of the population. Since a sample is expected to represent the whole population, the sampling procedure has to follow three fundamentals:

- 1. Should be representative.**
- 2. Large enough.**
- 3. The selected elements should have been properly approached, included and interviewed.**



REASONS FOR USING SAMPLES

There are many good reasons for studying a sample instead of an entire population:

- Samples can be studied more quickly than populations. Speed can be important if a physician needs to determine something quickly, such as a vaccine or treatment for a new disease.
- A study of a sample is less expensive than a study of an entire population because a smaller number of items or subjects are examined. This consideration is especially important in the design of large studies that require a long follow-up.
- A study of the entire populations is impossible in most situations.



STEPS IN SAMPLING

1. Definition of the population

We first need to identify the population we wish to draw the sample, from and do so somewhat formally because any **inferences** we draw are really only **applicable** to that population

2. Construction of a sampling frame (or thinking of an alternate)

The list of all possible units that might be drawn in a sample.



3. Selection of a sampling procedure

This is a critical decision about how to collect the sample. We will look at some different sampling procedure in the following slides.



TWO MAJOR TYPES OF SAMPLING PROCEDURES:

PROBABILITY

Each element has the same chance of being included in the sample.
Major types of probability sampling procedures:

1. **Simple random**
2. **Systematic**
3. **Cluster**
4. **Stratified**

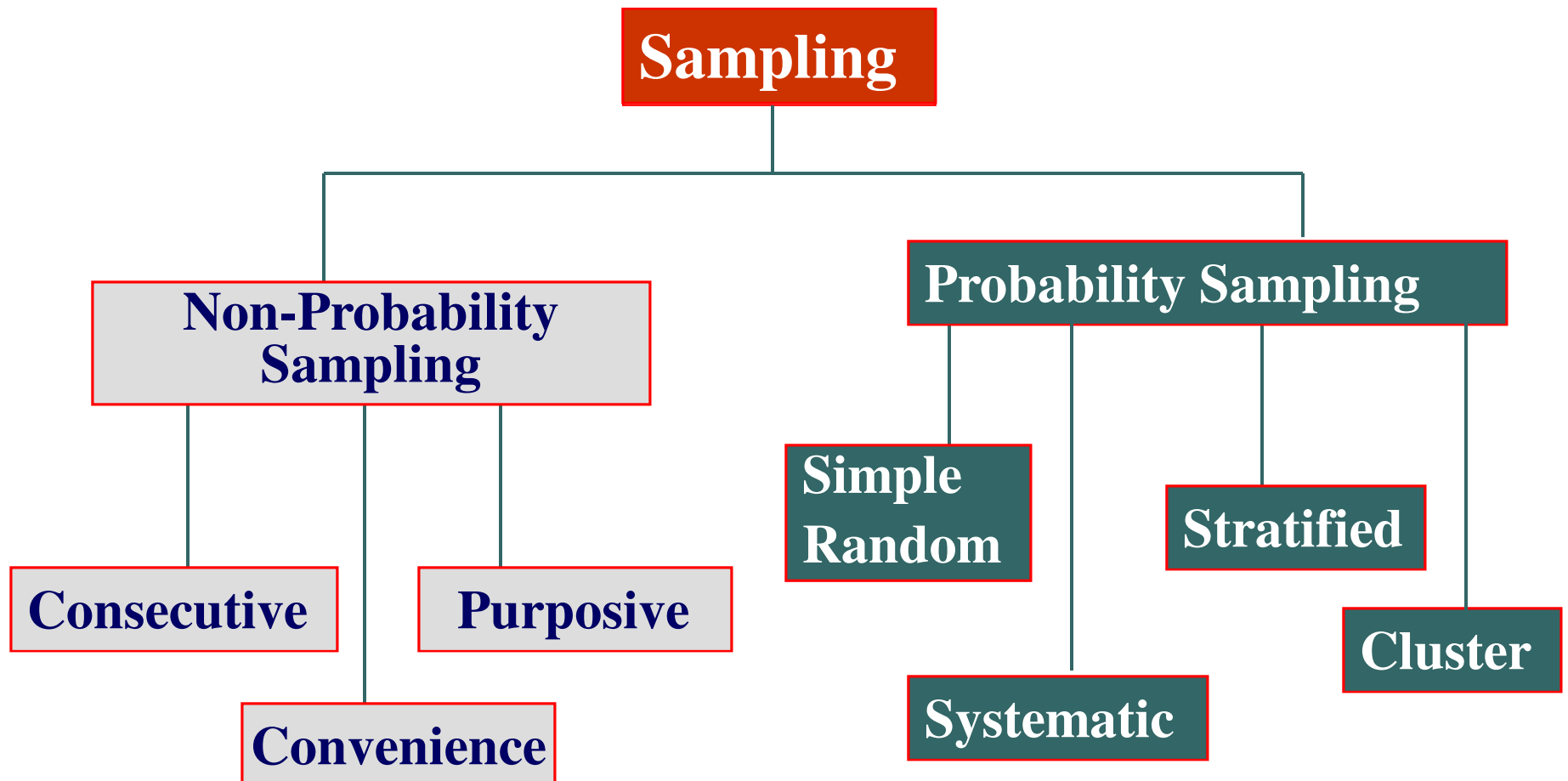
NON-PROBABILITY

There is no assurance that each element will have the same chance of being included in the sample. The 3 major types:

1. **Consecutive**
2. **Convenience**
3. **Purposive**



TYPES OF SAMPLING METHODS





SIMPLE RANDOM SAMPLING

PREREQUISITES

1. Sampling frame
a unique number is assigned to each **element**
2. Elements are selected into the sample randomly by 3 means
 - **Table of Random Numbers**
 - **Lottery Method**
 - **Computer Generated Numbers**



SYSTEMATIC SAMPLING

PREREQUISITES

1. Sampling frame (If available) if not then too systematic sampling can be undertaken. What is required is an estimate of population size and required sample size.



SYSTEMATIC SAMPLING

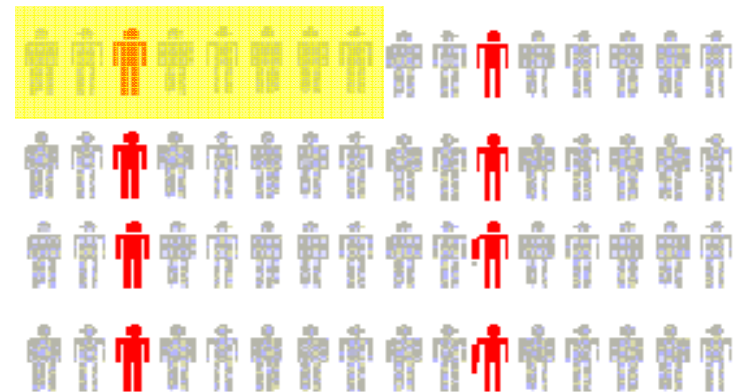
- Decide on sample size: n
- Determine population size = N
- Divide population of N individuals into groups of k individuals: $k = N/n$
- Randomly select one individual from the 1st group.
- Select every k -th individual thereafter.

$N = 640$

$n = 80$

$k = 8$

First Group





STRATIFIED SAMPLING

One of the main purposes of stratified sampling is to compare different strata, which may not be possible with simple random sampling alone.

Pre requisite: Sampling frame

- The population is first divided into groups of elements called strata.
- Each element in the population belongs to one and only one stratum.
- Best results are obtained when the elements within each stratum are as much alike as possible (i.e. homogeneous group).
- A simple random sample is taken from each stratum.



CLUSTER SAMPLING

When a list of the entire area is not available and it is not physically possible to visit the entire area (e.g. the city, or country) one can divide the area into several equal size clusters or units.

E.g.: Mohallas, Apartment Buildings, Villages, Schools

One can select (randomly) only a few cluster, number all the units within it and draw either:

1. A random sample or
2. A systematic sample



NONPROBABILITY SAMPLING

Nonprobability sampling design are often more practical than probability designs for some clinical research. Because statistical significance test are based on the assumption that a probability sample has been used, the objective in nonprobability sampling is to produce a facsimile, for the search question at hand of the probability sample.



Three major types of nonprobability sampling are

CONSECUTIVE

CONVENIENCE

PURPOSIVE



CONSECUTIVE SAMPLING

- It involves taking every patient who meets the selection criteria over a specified time interval or number of patients.
- It is the best of the nonprobability techniques and one that is very often practical.



CONVENIENCE SAMPLING

1. It is the process of taking those members of the accessible population who are easily available.
2. Sample is selected in a haphazard fashion.
3. It is widely used because of its obvious advantages in cost and logistics, however this type of sampling technique is fraught with biases.

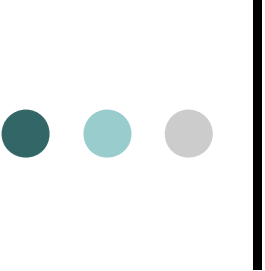


PURPOSIVE SAMPLING

- **Sampling is done on the basis of some pre determined idea (clinical knowledge etc).**
- **Specific targets are interviewed, because they posses the desired information.**

A decorative graphic consisting of three colored circles (dark teal, light teal, and grey) arranged horizontally, followed by a vertical black line.

VARIABLE



A Variable is a characteristic of a person, object or phenomenon that can take on different values.

A simple example of a variable is a person's age. The variable age can take on different values because a person can be 20 years old, 35 years old, and so on.



Dependent and independent variables

Because in health system research you often look for causal explanations, It is important to make distinction between dependent and independent variables.

The variable that is used to describe or measure the problem under study (outcome) is called the DEPENDENT variable.

The variables that are used to describe or measure the factors that are assumed to cause or at least to influence the problem are called the INDEPENDENT (exposure) variables.



Data

Data are values of the observation recorded for variables (e.g. age, weight, sex).



TYPES of DATA

Qualitative or categorical data:-

The characteristic which can't be expressed numerically like sex, ethnicity , healing etc.

Quantitative data or numerical data:-

The characteristic which can be expressed numerically like age, temperature, no. of children in a family.

Categorical Data

There are two types of categorical data:

- Nominal
- Ordinal data.



NOMINAL DATA

- In **NOMINAL DATA**, the variables are divided into named categories. These categories however, cannot be ordered one above another (as they are not greater or less than each other).

- *Example:*

<u>NOMINAL DATA</u>	<u>CATEGORIES</u>
---------------------	-------------------

Sex/ Gender:	male, female
Marital status:	single, married, widowed, separated, divorced



ORDINAL DATA

- In **ORDINAL DATA**, the variables are also divided into a number of categories, but they can be ordered one above another, from lowest to highest or vice versa.

- *Example:*

ORDINAL DATA

Level of knowledge:

Level of blood pressure:

CATEGORIES

good, average, poor

high, moderate, low



Quantitative Data

Interval Scale:-

You are also allowed to quantify the difference between two interval scale values but there is no natural zero.

Example :- Temperature scales are interval data with 25C warmer than 20C and a 5C difference has some physical meaning. Note that 0C is arbitrary, so that it does not make sense to say that 20C is twice as hot as 10C.



Quantitative data

- **Ratio Scale:-** You are also allowed to take ratios among ratio scaled variables. Physical measurements of height, weight, length are typically ratio variables. It is now meaningful to say that 10 m is twice as long as 5 m. This ratio hold true regardless of which scale the object is being measured in (e.g. meters or yards). This is because there is a natural zero.



Presentation of Data

- Data once collected should be presented in a such a way as to be easily understood . The style of presentation depends, of course, on type of data.
- Data can be presented in as frequency tables, charts, graphs, etc. Here we would discuss some of the important means of presentation.



FREQUENCY TABLES

- In a **FREQUENCY TABLE** data is presented in a tabular form. It gives the frequency with which (or the number of times) a particular value appears in the data.



Systolic Blood Pressure of patients coming to a tertiary care hospital OPD



n = 60

Distribution	Frequency	Relative	Cumulative Relative
Below 100	6	0.10	0.10
100 – 120	9	0.15	0.25
121 – 140	24	0.40	0.65
141 – 160	15	0.25	0.90
Above 160	6	0.10	1.00



Graphs

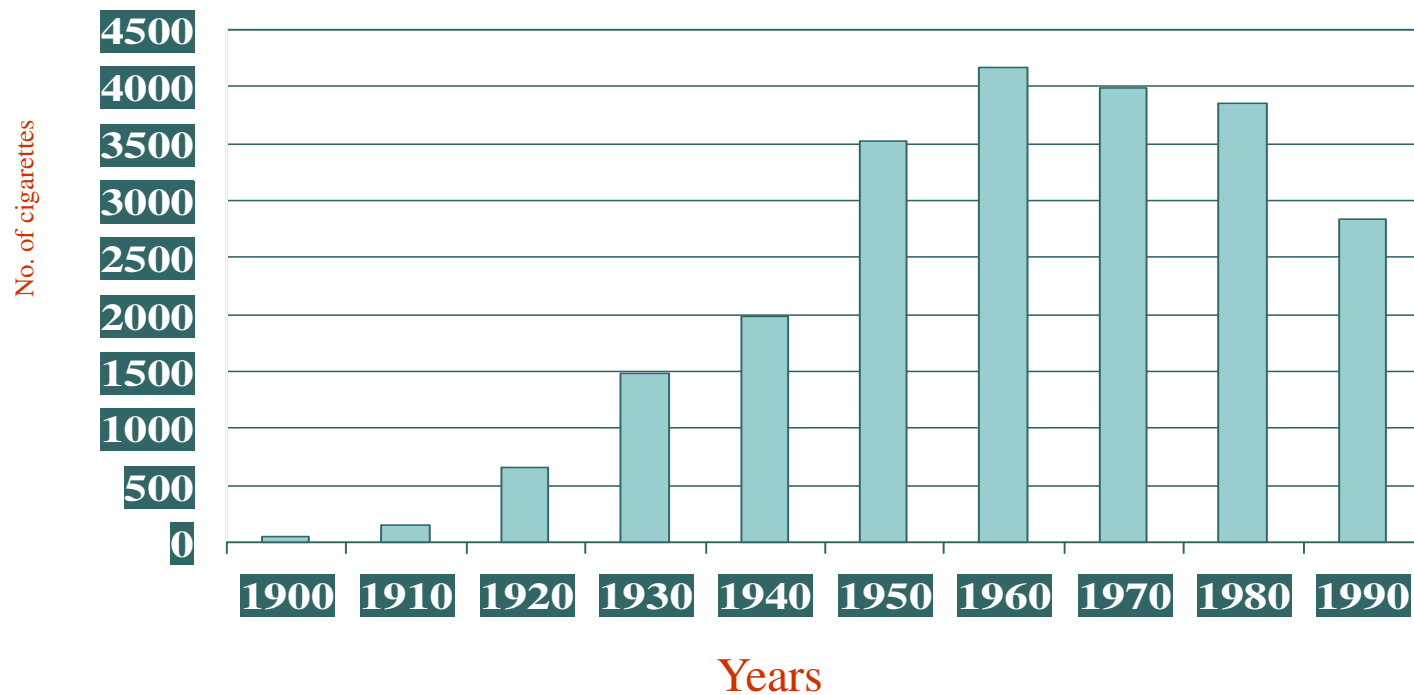
- Another way to summarize and display data is through the use of graph or pictorial representations of numerical data. Graphs should be designed so that they convey at a single glance the general patterns in a set of data.



Bar charts

- Bar charts are used for nominal or ordinal data.

Cigarette consumption of persons 18 years of age or older, United States, 1900 - 1990

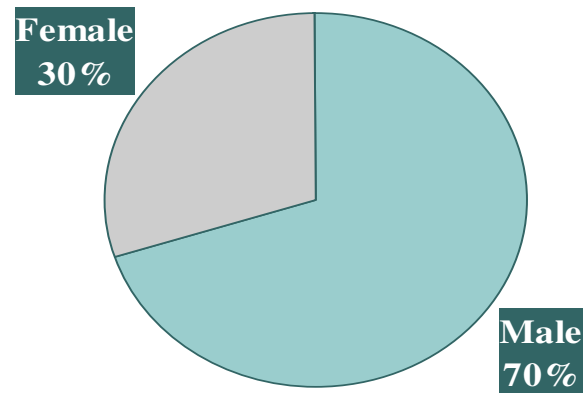




Pie chart

- Pie charts can also be used to display nominal or ordinal data.

Gender distribution

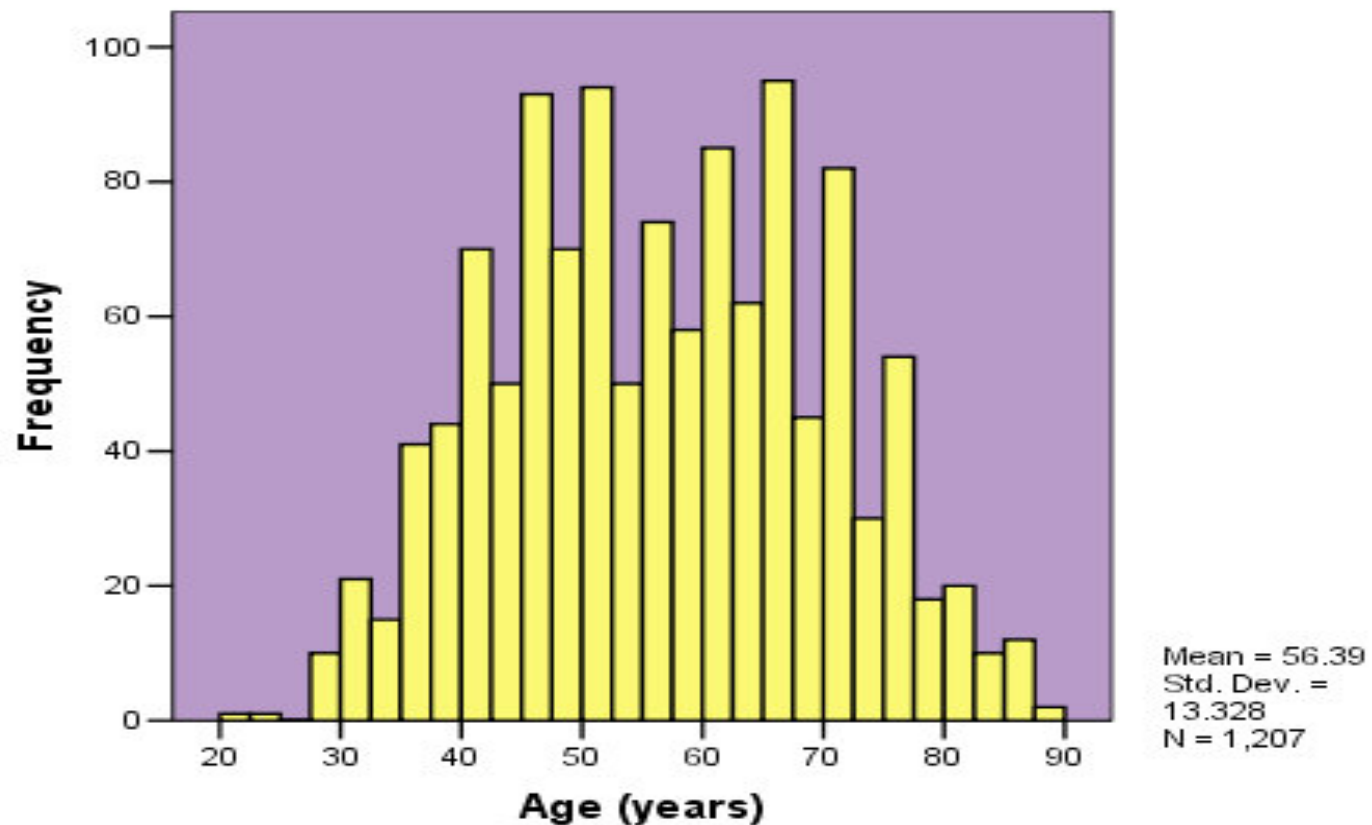




Histogram

A histogram depicts a frequency distribution for quantative data

Histogram showing distribution of Age (years)





SUMMARIZATION OF DATA



MEASURES OF CENTRAL TENDENCY

- *Mean*

- *Median*

- *Mode*



MEAN

- The **MEAN** (or **arithmetic mean**) is also known as the **AVERAGE**. It is calculated by totaling the results of all the observations and dividing by the total number of observations. Note that the mean can only be calculated for numerical data.



● ● ● | **MEDIAN**

The MEDIAN is the value that divides a distribution into two equal halves.

- The median is useful when some measurements are much bigger or much smaller than the rest. The mean of such data will be biased toward these extreme values.
- The median is not influenced by extreme values.



MODE

- The **MODE** is the most frequently occurring value in a set of observations.



MEASURES OF VARIATION

Range is defined as the difference in value between the highest (maximum) and the lowest (minimum) observation

Variance Quantifies the amount of variability or spread about the mean of the sample.

Standard deviation it is the square root of the variance



Standard Deviation

- The **STANDARD DEVIATION** is a measure, which describes how much individual measurements differ, on the average, from the mean.
- A large standard deviation shows that there is a wide scatter of measured values around the mean, while a small standard deviation shows that the individual values are concentrated around the mean with little variation among them.



Standard error of the mean

When we draw a sample from study population and compute its sample mean it is not likely to be identical to the population mean. If we draw another sample from same population and compute its sample mean, this may also not be identical to the first sample mean. It probably also differs from the true mean of the total population from which the sample was drawn this phenomenon is called sampling variation.

Standard error:- The standard error gives an estimate of the degree to which the sample mean varies from the population mean and this measures is used to calculate CI.



THE NORMAL DISTRIBUTION

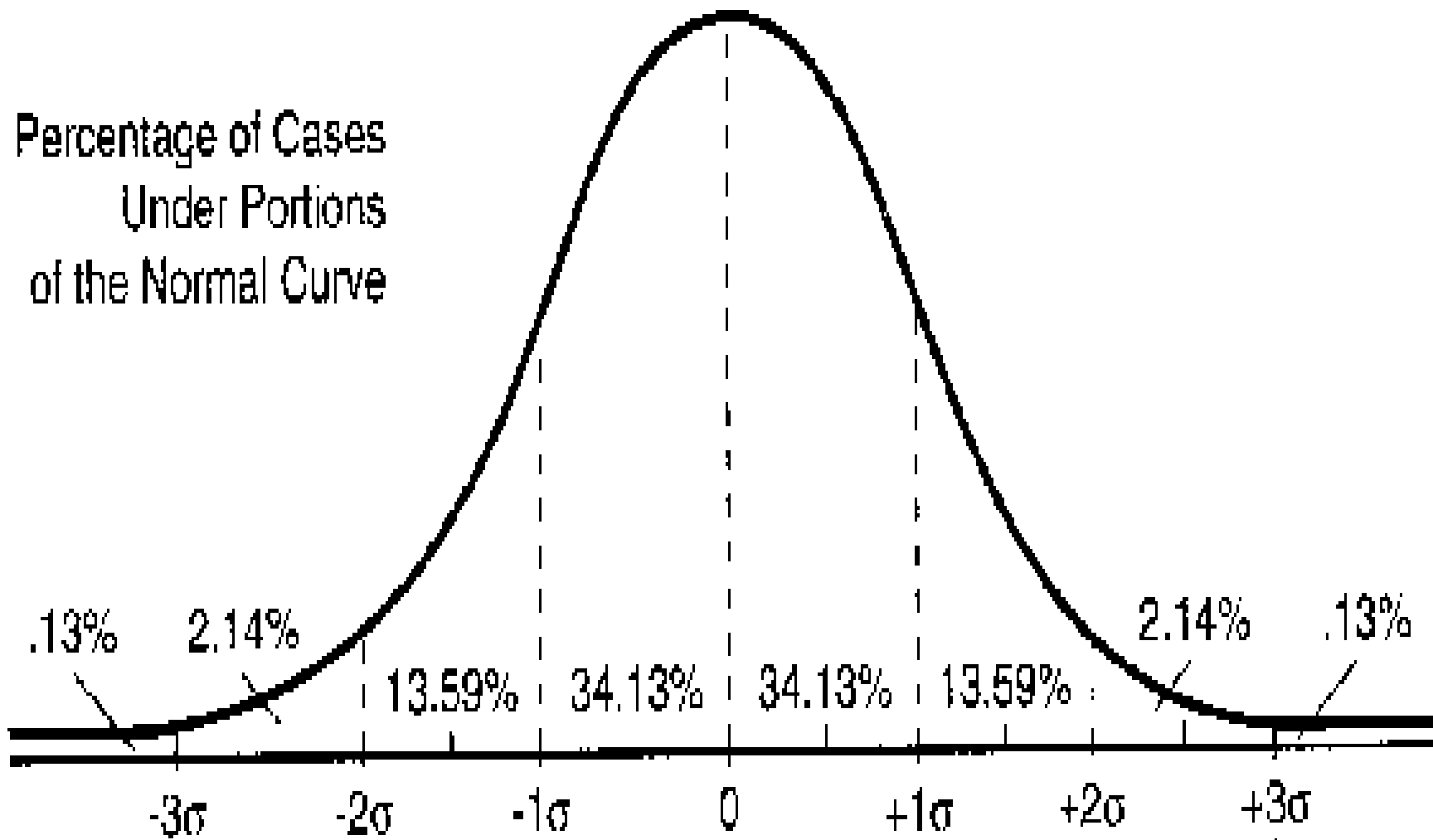
- Many variables have a normal distribution. This is a bell shaped curve with most of the values clustered near the mean and a few values out near the tails.



- The normal distribution is symmetrical around the mean. The mean, the median and the mode of a normal distribution have the same value.
- An important characteristic of a normally distributed variable is that 95% of the measurements have value which are approximately within 2 standard deviations (SD) of the mean.



THE NORMAL DISTRIBUTION





Estimation

The process of using sample information to draw conclusion about the value of a population parameter is known as estimation.



Point Estimate

- A point estimate is a specific numerical value estimate of a parameter.
- The best point estimate of the population mean μ is the sample mean
 \bar{X}
- But how good is a point estimate?
- There is no way of knowing how close the point estimate is to the population mean
- Statisticians prefer another type of estimate called an interval estimate



Interval Estimate

- An interval estimate of a parameter is an interval or a range of values used to estimate the parameter

Confidence Level

- The confidence level of an interval estimate of a parameter is *the probability that the interval estimate will contain the parameter*
- Three commonly used confidence levels are 90%, 95% and 99%
- If one desires to be more confident then the sample size must be larger



MEASURES OF DISEASE FREQUENCY

A Presentation



RATIO

- The most basic measure of distribution.
- Obtained by simply dividing one quantity by another without implying any specific relationship between the numerator and denominator, such as the number of stillbirths per thousand live births.
- In ratio, the numerator & denominator are mutually exclusive.



PROPORTION

- A proportion is a type of ratio in which those who are included in the numerator must also be included in the denominator.
- For example: the proportion of women over the age of 50 who have had a hysterectomy, or the number of fetal deaths out of the total number of births (live births plus fetal deaths).



RATE

- A rate is a proportion with specifications of time. There is a distinct relationship between the numerator and denominator with a measure of time being an intrinsic part of the denominator.
- For example, the number of newly diagnosed cases of breast cancer per 100,000 women during a given year.



IMPORTANT POINT

- It is necessary to be very specific about what constitutes both the numerator and the denominator. In some circumstances, it is important to make clear whether the measure represents the number of events or the number of individuals.
- For example, the frequency of myopia among a population of school children could represent the number of affected eyes in relation to total eyes, or the number of children affected in one or both eyes relative to all students.



PREVALENCE

- Prevalence quantifies the proportion of individuals in a population who have the disease at a specific instant and provides an estimate of the probability (risk) that an individual will be ill at a point in time
- The *formula for calculating the prevalence* $P = \text{number of existing cases of a disease} / \text{total population (at a given point in time)}$



POINT PREVALENCE

- Prevalence can be thought of as the status of the disease in a population at a point in time and as such is also referred to as point prevalence.
- This "point" can refer to a specific point in calendar time or to a fixed point in the course of events that varies in real time from person to person, such as the onset of menopause or puberty or the third postoperative day.



PERIOD PREVALENCE

- It represents the proportion of cases that exist within a population at any point during a specified period of time.
- The numerator thus includes cases that were present at the start of the period plus new cases that developed during this time

E.g. Frequency of patients receiving Psychiatric Rx between May 31 – Dec 01 2008



INCIDENCE:

- Incidence quantifies the number of new events or cases of disease that develop in a population of individuals at risk during a specified time interval.



Cumulative incidence (CI)

- Is the proportion of people who become diseased during a specified period of time. It provides an estimate of the probability, or risk, that an individual will develop a disease during a specified period of time

$$CI = \frac{\text{No. of new cases of a disease}}{\text{Total population at risk}}$$



Issues in the Calculation of Measures of Incidence

- For any measure of disease frequency, precise definition of the denominator is essential for both accuracy and clarity. This is a particular concern in the calculation of incidence. The denominator of a measure of incidence should include only those who are considered "at risk" of developing the disease.



Contd.

- That is, the total population from which the new cases could arise.
Consequently, those who currently have or have already had the disease under study or persons who cannot develop the disease for reasons such as age, immunization, or prior removal of the involved organ should be excluded from the denominator.



Special Types of Incidence Rates

MORBIDITY RATE

Is the incidence rate of non fatal cases in the total population at risk during a specified period of time.

For example, the morbidity rate of tuberculosis (TB) in the U.S. in 1982 can be calculated by dividing the number of nonfatal cases newly reported during that year by the total U.S. midyear population.

Total no of nonfatal cases of TB in POP at risk

Mid year POP

25,520

231,534,000

= 11.0 per 100,000 population



MORTALITY RATE

- It expresses the incidence of deaths in a particular population during a period of time.
- It is calculated by dividing the number of fatalities during that period by the total population.
- This can be further divided into cause specific or all case mortality.



Measures of Association



Measures of Association

- Relative risk (cohort study)
- Odds ratio (case control)



Cohort Studies

	<i>Diseased</i>	<i>Non Diseased</i>	
Exposed	a	b	a+b
Non Exposed	c	d	c+d



Relative Risk

- Incidence in exposed individuals = $a/a+b$
Or proportion of exposed people who developed the disease
- Incidence in non-exposed individuals = $c/c+d$
Or proportion of non exposed people who develop disease

Relative Risk = $\frac{\text{Incidence in exposed}}{\text{Incidence in non exposed}}$

RR = $\frac{a/a+b}{c/c+d}$



Calculating the Relative Risk

	Disease Status		
	CHD +	CHD -	Total
Smoker	112	176	288
Non smoker	88	224	312

Incidence in exposed = $a / a+b = 112 / 288 = 0.38$

Incidence in non exposed = $c / c+d = 88 / 312 = 0.28$

$RR = 0.38 / 0.28 = 1.38$



Interpretation of RR

- Compared to non smokers, the smokers have a 1.38 times greater risk of developing CHD



Odds Ratio

- Incidence cannot be measured in case control studies because we start with the diseased people (cases) and non diseased people (controls), hence we calculate OR



Case Control

	Cases	Controls	
<i>Exposed</i>	a	b	a+b
Non Exposed	c	d	c+d
	a+c	b+d	

$$OR = a/c \div b/d \text{ or } ad/bc$$



Passive Smoking & Breast Cancer

Exposed (Passive
Smokers)

Not exposed

Breast cancer	No Breast cancer	Total
140 (a)	370 (b)	510
40 (c)	234 (d)	274

$$\text{Odds} = 140 / 40 = 3.5$$

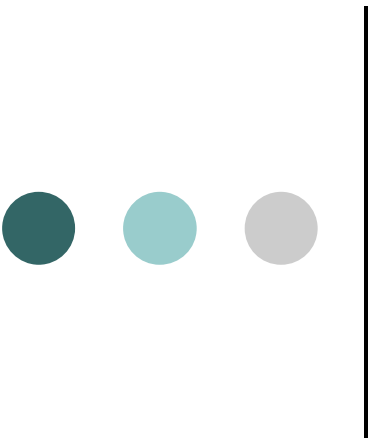
$$\text{Odds} = 370 / 234 = 1.6$$

$$\text{OR} = 3.5 / 1.6 = 2.2$$

Compared to the control, the odds of being a passive smoker are 2.2 > in Ca breast cases



Factors Affecting Study Outcomes



BIAS



Bias is:

Any systematic error that results in an incorrect estimate of the association between the exposure and outcome. Usually introduced by the experimenter or the researcher himself due to non-standardized measuring techniques.



Type of Bias:-

Selection Bias

**Observation (Information/Misclassification)
Bias**

Recall Bias

Interviewers Bias

Lost-to-follow up



Can Control Bias:- In study design through

Choice of study population

Data collection:-

Uniform Source of information

Efficient Questionnaire development

**Standardization of measurement
technique**

Blinding



Confounding



The concept of confounding is a central one in the interpretation of any epidemiological study.

Confounding can be thought of as mixing of the effect of the exposure under study on the disease with that of an extraneous factor.

This external factor or variable must be associated with the exposure and, independent of the exposure must be a risk factor for the disease.



Example of confounding





Table 1. Relation of Myocardial infarction (MI) to Recent Oral Contraceptive (OC) Use.

	MI +ve	MI -ve	Estimated relative risk
OC			
Yes	29	135	=1.68
No	205	1607	
Total	234	1742	



Table2:-Age -specific Relation of Myocardial infarction (MI) to recent Oral Contraceptive (OC) Use.

Age (yrs)	Recent OC use	MI +ve	MI -ve	Estimated age-Specific relative risk
25 – 29	Yes	4	62	7.2
	No	2	224	
30 – 34	Yes	9	33	8.9
	No	12	390	
35 – 39	Yes	4	26	1.5
	No	33	330	
40 – 44	Yes	6	9	3.7
	No	65	362	
45 – 49	Yes	6	5	3.9
	No	93	301	
Total		234	1742	



Confounding can be controlled in study design through:

- Restriction
 - Matching exposure
- Randomization

Confounding can be controlled in analysis through:

- Stratification
- Multivariate analysis



The **role of confounding, chance and bias** have to be **evaluated** in studies appropriate selection of the population to be studied , with proper study design, so that the **results** can be **applied** to **other population** i.e., they are **valid** and **generalizable**.



● ● ● | **Evaluation of the role of chance consists of two components:-**

1. **Hypothesis testing**
2. ***Estimation of the confidence interval***



FORMULATION & TESTING OF HYPOTHESIS



WHAT IS HYPOTHESIS?

Hypothesis: A testable theory, or statement of belief used in evaluation of a population parameter of interest e.g. Mean or proportion



- Suppose a study is being conducted to answer questions about differences between two regimens for the management of diarrhea in children:
the sugar based modern ORS and the time-tested indigenous herbal solution made from locally available herbs.
- One question that could be asked is:
"In the population is there a difference in overall improvement (after three days of treatment) between the ORS and the herbal solution?"



- There could be only two answers to this question:
- Yes
- No



Null Hypothesis

"There is no difference between the 2 regimens in term of improvement" (null hypothesis).

A null hypothesis is usually a statement that there is no difference between groups or that one factor is not dependent on another and corresponds to the No answer.



Alternative Hypothesis

- "There is a difference in terms of improvement achieved by a three days treatment with the ORS and that of the herbal solution" (alternative hypothesis).
- Associated with the null hypothesis there is always another hypothesis or implied statement concerning the true relationship among the variables or conditions under study if no is an implausible answer. This statement is called the alternative hypothesis and corresponds to the "Yes" answer.



TYPES OF ALTERNATE HYPOTHESIS

- o **Directional**
- o **Non Directional**



WHY TEST HYPOTHESIS

Hypothesis testing permits generalization of an association or a difference obtained from a sample to the population from which it came.

Hypothesis testing involves conducting a test of statistical significance and quantifying the degree to which sampling variability may account for the result observed in a particular study. It entails the following steps.



STEPS IN HYPOTHESIS TESTING

1. Statement of research question in terms of statistical hypothesis (Null and alternate hypothesis)
2. Selection of an appropriate level of significance. The significance level is the risk we are willing to take that a sample which showed a difference was misleading. 5% significance level means that we are ready to take a 5% chance of wrong results.



STEPS IN HYPOTHESIS TESTING

3. Choosing an appropriate statistics
 t test, z test for continuous data, chi square for proportions etc.

Test statistics is computed from the sample data and is used to determine whether the null hypothesis should be rejected or retained.

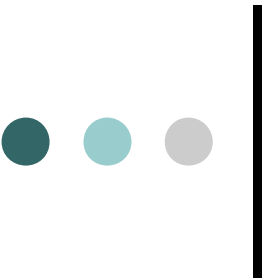
Test statistics generates p value



***P* value:** Indicates the probability or likelihood of obtaining a result at least as extreme as that observed in a study by chance alone, assuming that there is truly no association between exposure and outcome under consideration.

By convention the *p* value is set at 0.05 level. Thus any value of *p* less than or equal to 0.05 indicates that there is at most a 5% probability of observing an association as large or larger than that found in the study due to chance alone given that there is no association between exposure and outcome. If *p* value > 0.05 do not reject the null hypothesis .

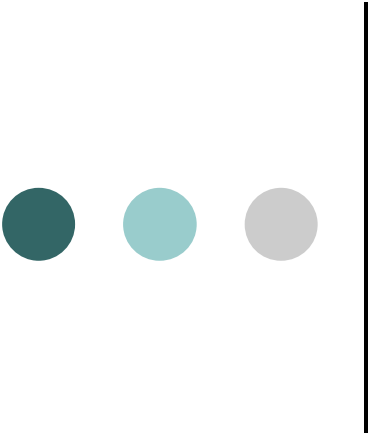


- 
4. Performing calculations and obtaining p value
 5. Drawing conclusions, rejecting null hypothesis if the p value is less than the set significance level



α and β error

Tests of Significance	True Ho Hypothesis	False Ho Hypothesis
Accept Ho Hypothesis	Correct Decision ✓	Wrong Decision β Error
Reject Ho Hypothesis	Wrong Decision α Error	Correct Decision ✓



SAMPLE SIZE ESTIMATION



Sample size calculations depend on:

1. Type of study.
2. Magnitude of the outcome of interest derived from previous studies.
3. Type of statistical analysis required (comparing means or proportions).
4. Level of significance / Power.



Sample size for single proportion depends on:

- 1. The prevalence of the condition/attribute of interest.**
- 2. Level of confidence.**
- 3. Margin of error.**



Example of Sample size calculation for single proportion



- A local health department wishes to estimate the prevalence of tuberculosis among children under 5 year of age in a locality. How many children should be included in the sample so that the prevalence may be estimated within 5% point of the true value with 95% confidence, if it is known that the true rate is unlikely to exceed 20%?

Sample size calculation and formula for single proportion



Perform Estimation

1.1. Estimating a population proportion with specified absolute precision

Please select the desired unknown:

- ☐ Confidence level (%)
- ☐ Anticipated population proportion
- ☐ Absolute precision required
- ☒ Sample size

Please enter the remaining values:

$1 - \alpha$ 95

P 0.20

d 0.05

n 246

$$n = \frac{z_{1-\alpha/2}^2 P(1 - P)}{d^2}$$

Print

Help

Close



Sample size for single group mean depends on:

- 1. The Mean of the condition of interest.**
- 2. Level of confidence.**
- 3. Margin of error.**



Example of Sample size calculation for single group mean

- A district medical officer seeks to estimate the mean hemoglobin level among pregnant women in his district. A previous study of pregnant women showed average hemoglobin level 8.2 g/dl and standard deviation of 4.2 g/dl. Assuming a sample of pregnant women is to be selected, how many pregnant women must be studied if he wanted the estimate should fall within 1 g/dl with 95% confidence?

Sample size calculation and formula for single group mean



Perform Estimation

7.1. Estimating a population mean

Please select the desired unknown:

- ☐ Confidence level (%)
- ☐ Absolute precision required
- ☐ Relative precision
- ☐ Population mean
- ☐ Population standard deviation
- ☐ Population variance
- ☒ Sample size

Please enter the remaining values:

$1 - \alpha$	95
d	1
ε	0.12195121951
μ	8.2
σ	4.2
σ^2	17.64
n	68

$$n = \frac{z_{1-\alpha/2}^2 \sigma^2}{d^2} \text{ or } \frac{z_{1-\alpha/2}^2 \sigma^2}{\varepsilon^2 \mu^2}$$

Print Help Close

Where $\varepsilon = d/\mu$



Sample size for two proportions depends on:

- 1. The prevalence of the condition /
attribute of interest for both groups.**
- 2. Level of confidence.**
- 3. Power of the test.**



Example of Sample size calculation for two proportions

- It is believed that the proportion of patient who develop complications after undergoing one type of surgery is 5% while the proportion of the patients who develop complication after a second type of surgery is 15%. How large should the sample size be in each of the two groups of patients if an investigator wishes to detect with a power of 90%, whether the second procedure has a complication rate significantly higher than the first at the 5% level of significance?

Sample size calculation and formula for two proportions



Perform Estimation

2.2a. Hypothesis tests for two population proportions (one-sided test)

Please select the desired unknown:

- ☐ Level of significance (%)
- ☐ Power of the test (%)
- ☐ Anticipated population proportion 1
- ☐ Anticipated population proportion 2
- ☒ Sample size

Please enter the remaining values:

α 5

$1 - \beta$ 90

P_1 0.05

P_2 0.15

n 153

$$n = \frac{\left\{ z_{1-\alpha} \sqrt{2\bar{P}(1-\bar{P})} + z_{1-\beta} \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right\}^2}{(P_1 - P_2)^2}$$

Print

Help

Close



Sample size for two group means depends on:

- 1. The means/variance for both groups.**
- 2. Level of confidence.**
- 3. Power of the test.**



Example of Sample size calculation for two group means



Suppose the true mean systolic blood pressure (SBP) of 35 to 39 year old OC users is (132.86 mmHg) and standard deviation (15.34 mmHg). Similarly, for non-OC users, the mean SBP is (127.44 mmHg) with standard deviation (18.23 mmHg). If we desire to estimate the difference between 2 groups of equal size, what would be the minimal sample size required with a power of 80% at 95% confidence level?



Calculator

Sample Size For Comparing Two Means

Confidence Interval % (two-sided)	95	Enter a value between 0 and 100, usually 95%		
Power	80	Enter a value between 0 and 100, usually 80%		
Ratio of sample size (Group 2/Group 1)	1			
	Group 1		Group 2	Enter means OR difference on next line
Mean	132.86	and	127.44	or Difference
Std. Dev.	15.34		18.23	Enter Std. Deviation OR Variance of each individual group
Variance				



Sample size - Calculation

Sample Size For Comparing Two Means

Input Data

Confidence Interval (2-sided)	95%
Power	80%
Ratio of sample size (Group 2/Group 1)	1

	Group 1	Group 2	Mean difference ¹
Mean	132.86	127.44	5.42
Standard deviation	15.34	18.23	
Variance	235.316	332.333	

Sample size of Group 1	152
Sample size of Group 2	152
Total sample size	304



● ● ● | **Sample size for sensitivity and specificity depends on:**

- 1. The prevalence of the condition/attribute of interest.**
- 2. Estimated sensitivity.**
- 3. Estimated specificity.**
- 4. Level of significance.**
- 5. Margin of error.**



Example of Sample size calculation for sensitivity and specificity

- If we want to determine the sensitivity and specificity of graded compression ultrasonography in the diagnosis of acute appendicitis by the gold standard histopathology. How many patients should be included in the sample .The prevalence OF AA is 77% and estimated sensitivity of US is 96.5% and estimated specificity is 94.1% with 95% confidence, if we want to keep margin of error as 10%?

Sample size calculation and formula for sensitivity and specificity studies



Sample size calculation for Sensitivity & Specificity Studies

Instruction:

Enter five values in yellow areas only. Then, read the sample size in blue colour box.

If you want to know how it is calculated, read the explanation in the next sheet.

Expected Sensitivity	0.97	◀ From literature or pilot study	
Expected Specificity	0.94	◀ From literature or pilot study	
Expected Prevalence	0.77	◀ From literature or pilot study	
Desired Precision	0.10	◀ Researcher's judgment	The precision is too big for both Sensitivity and Specificity. Try a smaller one.
Confidence level	95%	◀ 95% is recommended.	

To achieve the precision of 0.1 for 'Sensitivity', we need 'the total sample size' of

= 17 ◀ Warning: This sample size will not give the precision of 0.1 for Specificity.

With this sample size, the precision for 'Specificity' will be

= 0.234

To achieve the precision of 0.1 for 'Specificity', we need 'the total sample size' of

= 96 ◀ This is preferable as it will give the precision of 0.1 or less for both Sensitivity and Specificity.

With this sample size, the precision for 'Sensitivity' will be

= 0.042

Written by Dr Lin Naing @ Mohd. Ayub Sadiq, School of Dental Sciences, Universiti Sains Malaysia (naing@kck.usm.my)

dated 23-Mar-2004



Suggested websites for sample size calculators

1. <http://www.raosoft.com/samplesize.html>
2. <http://www.quantitativeskills.com/sisa/calculations/samsize.htm>
3. <http://www.openepi.com/Menu/OpenEpiMenu.htm>



Screening



Screening

- Screening for disease control can be defined as the examination of asymptomatic people in order to classify them as likely or unlikely to have the disease that is object of screening.
- If done in large groups---mass screening or population screening.



Characteristics of Disease to be Screened

- Disease must pass through preclinical phase during which it is undiagnosed but detectable
- Early treatment must offer some advantage



Validity

- The ability of a test to distinguish between who has disease and who does not



Sensitivity

- Of a test is its ability to detect people who do have disease.
- If a Test is always positive for all diseased persons then sensitivity of the Test will be 100%.



Specificity

- It is the ability of a Test to detect people who don't have disease.
- Thus a Test which is always negative in non- diseased individuals is called to have 100% specificity.



Validity



	Diseased	Non Diseased	
Positive	<div>TP</div> <div>a</div>	<div>FP</div> <div>b</div>	a+b
Negative	<div>FN</div> <div>c</div>	<div>TN</div> <div>d</div>	c+d



Test FNA	CA Breast Positive	CA Breast Negative	Total
Positive	60 a + +	50 b - +	a + b 110
Negative	20 c - +	70 d - -	c +d 90
Total	80 a + c	120 b + d	a + b + c + d 200



- Sensitivity = $\frac{a}{a + c} \times 100 = \frac{60}{80} \times 100 = 75\%$

I.e. Test (FNA) is 75% sensitive in detecting disease

- Specificity = $\frac{d}{d + b} \times 100 = \frac{70}{120} \times 100 = 58\%$

I.e. Specificity of (FNA) is 58% to detect non-diseased persons



Positive Predictive Value i.e PPV

$$\text{PPV} = \frac{a}{a + b} \times 100 = \frac{60}{110} \times 100 = 55\%$$

I.e. 55% persons are actually suffering from disease.

$PPV \propto \text{Prevalence}$

Negative Predictive Value i.e NPV

$$\text{NPV} = \frac{d}{c + d} \times 100 = \frac{70}{90} \times 100 = 78\%$$

I.e. 78% persons are actually free from disease.



Test	Disease Present	Disease Not Present	Total
Positive	True Positive (TP) + +	False Positive (FP) - +	TP + FP
Negative	False Negative (FN) + -	True Negative (TN) - -	FN + TN
Total	TP + FN	TN + FP	TP+FP+TN+FN

•Sensitivity = $\frac{TP}{TP + FN} \times 100$

•Specificity = $\frac{TN}{TN + FP} \times 100$

190

•PPV= $\frac{TP}{TP + FP} \times 100$

•NPV = $\frac{TN}{TN + FN} \times 100$



Relationship of Disease Prevalence to PPV

Example: Sensitivity = 99%; Specificity = 95% In a population of 10,000 with a disease prevalence of 1%

Dis. Prev	Test Results	Disease	Not Disease	Total
1%	Positive	99	495	594
	Negative	1	9405	9406
	Total	100	9900	10,000

$$\text{PPV} = 99/594 = 17\%$$



Relationship of Disease Prevalence to PPV

Example: Sensitivity = 99%; Specificity = 95% In a population of 10,000 with a disease prevalence of 5%

Dis. Prev	Test Results	Disease	Not Disease	Total
5%	Positive	495	475	970
	Negative	5	9025	9030
	Total	500	9500	10,000

$$\text{PPV} = 495 / 970 = 51\%$$



Relationship between PPV & Prevalence

- A screening program is most effective and beneficial if it is directed to a high-risk target population
- Screening a total population for a relatively infrequent disease can be very wasteful of resources and may yield very few previously undetected cases



DEVELOPING THE QUESTIONNAIRE



POINTS OF IMPORT IN DESIGNING A QUESTIONNAIRE

- It should be ensured that the format of the questionnaire be attractive and easy for the respondents to fill, overcrowding or clutter should be avoided and all questions and pages clearly numbered
- The questionnaire should not be too long
- To maintain flow of the instrument, questions concerning major areas should be grouped together
- Simple questions about age, birth date etc should be put at the beginning to warm up the respondent

POINTS OF IMPORT IN DESIGNING A QUESTIONNAIRE



- Questions should be close ended, possible answers to close ended questions should be lined vertically, preceded by boxes, brackets or numbers

Example

How many different medicines do you take daily (check one)

☐ None

☐ 1-2

☐ 3-4

☐ 5-6

☐ 7 or more



POINTS OF IMPORT IN DESIGNING A QUESTIONNAIRE

- If more details are required pertaining to a question , then the filter/skip technique should be used to save time and allow respondents to avoid irrelevant questions.

Example :Have you ever been told that you have hypertension.

Yes

No

If yes proceed to next question

How long back were you told that you have hypertension



POINTS OF IMPORT IN DESIGNING A QUESTIONNAIRE



- Wording of questions should be simple and free from ambiguity, non judgmental and be soliciting only one response.
- For behaviors that may change overtime specific time span should be asked for in the question

Example :During the past 12 months how many doctor visits did you make.

- Always choose a appropriate means of measurement e.g. score /scales.



POINTS OF IMPORT IN DESIGNING A QUESTIONNAIRE

- Sensitive topic questions should be left for the end
- If similar research instruments are available it may be a good idea to review and if required borrow questions.
- Always try to ensure that if questions are to be asked in any language besides English they shall be so written too

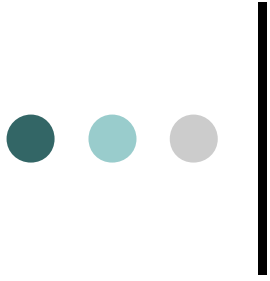
A decorative graphic consisting of three colored circles (dark teal, light teal, and grey) arranged horizontally, followed by a vertical black line.

Writing a Synopsis



Introduction

- Synopsis: A short account of something longer. A document that contains a brief information about the research a postgraduate trainee plans to conduct.



1. Topic

The title of synopsis / research protocol precisely reflects the objective (s) of the proposed study in self-explanatory words in one or two lines.



2. Introduction

- Magnitude of the problem indicated clearly
- Recent (last 5 years) local & international knowledge cited/provided (if not available, that stated too)
- Gaps identified in existing knowledge regarding the specific topic/theme of study, what would the current study add to the same
- Rationale of study led by the above stated two points



3. Objective(s):

Should be stated in clear & quantifiable words

4. Operational definition:

Should clearly define and indicate means of measurement/determination of all exposure/outcome (dependent & independent) variables indicated in the objectives

5. Hypothesis:

Appropriately framed in terms of objectives (alternate hypothesis)



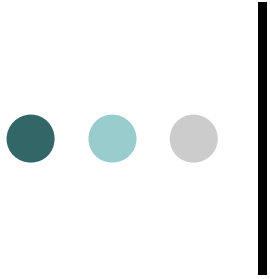
4. Material and Methods:

- Setting: Place of study e.g. unit / department of the institute
- Duration of study: Time period required to complete study (should not be less than 6 months) & to be commenced after the approval of synopsis
- Sample Size: Sample size calculation done or stated with justification, based on previous literature review
- Sample Selection:
Inclusion and exclusion criteria described with justification



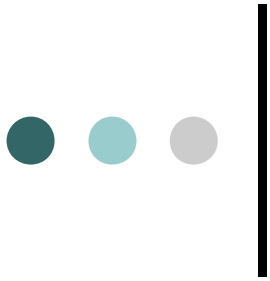
4. Material and Methods:

- Study Design: Name of appropriate study design stated
- Data Collection:
 - Description of study population & methods of enrolment
 - Selection & means of obtaining information of variables of interest clearly indicated
 - Means of determination of main outcomes indicated, by who & when
 - Means of control of bias & confounding indicated



Data Collection (cont.)

- Sources of variables / data collection clearly stated; if several methods were used, like records, interview, laboratory tests, other measurements like anthropometry, blood pressure etc., then all stated in reference to time and source
- Means of control of confounding variables (if to be done at analysis stage) indicated



Data Collection (cont.)

- Ethical issues considered and maintained through out the study
- Type of informed consent taken
- Proforma attached in appropriate format, indicating all variables stated in methods and relevant to the objective and not in the format of a history sheet



Data Analysis:

- Statistical software used for data entry stated
- Statistical software used for data analyses stated
- Plan of data analyses according to objectives, study design, and methods
- Type of descriptive statistics and distribution of variables stated



Data Analysis:

- Any new variables that could be formed defined and stated e.g. a scoring system, body mass index , % body fat etc. If composite variable formed then the same clarified & indicated in operational definition too
- Statistical tests used for testing of hypothesis with levels of significance
- Any method used for control of confounding factors stated / described



References:

- Minimum of FIVE
- References cited in Vancouver style
- All references could be verified

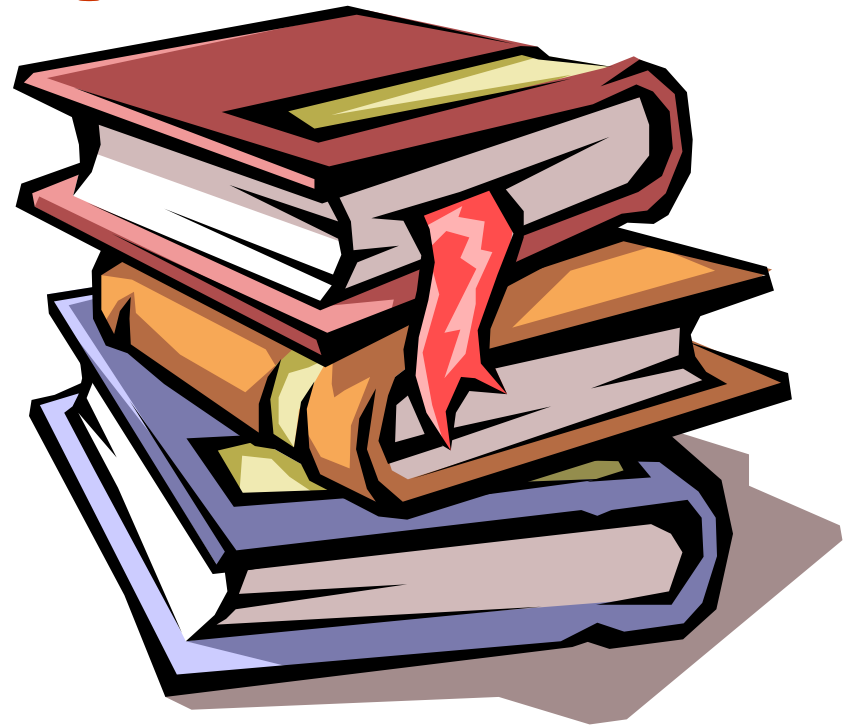


DISSERTATION WRITING

WHAT?

&

WHY?





Dissertation Writing



Detailed discourse on a subject especially submitted for a higher degree in a University [Oxford]

STEPS IN WRITING A DISSERTATION



Application for Training Registration
(immediately on start of training)

Training Registration

Approval of Supervisor

Submission of Synopsis
(within 6 months of registration)

Approval of Synopsis by the Synopsis Review Committee of CPSP

Dissertation writing & its Submission
(at least 6 months before the examination in which appearing)

Assessment of Dissertation by External Assessors

Approval of Dissertation

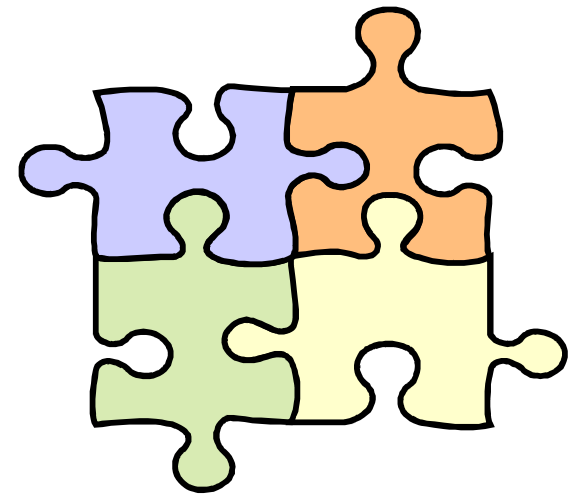
Eligibility for FCPS II Examination



FORMAT OF DISSERTATION

Part - 1

- Title page
- Supervisor's certificate
- Dedication
- Acknowledgement
- Table of contents
- List of tables
- List of figures, graphs, illustrations
- List of abbreviations

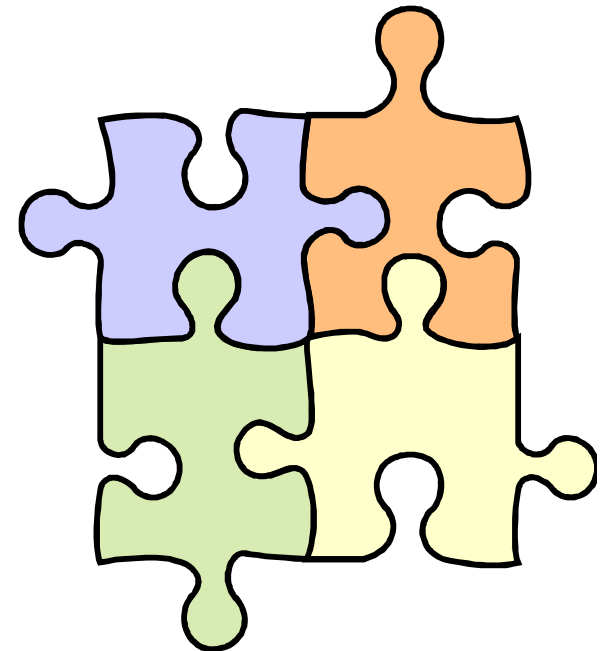




FORMAT (contd.)

Part - 2 (about 70-100 pages)

- Abstract
- Introduction
- Review of literature
- Objective(s) of study
- Operational definitions
- Hypothesis
- Material & Methods
- Result
- Discussion
- Conclusion(s)
- References (Bibliography)
- Annexes (Proforma etc.)





Title Page (Sample)

EDUCATIONAL IMPACT OF IMPROVED MONITORING OF CPSP FELLOWSHIP TRAINING PROGRAMS

by

Dr. Syeda K. Ali, MHPE

Supervisor:

Prof. Neil Paget, Ph.D

Work done at

Department of Medical Education

College of Physicians & Surgeons Pakistan

JUNE 09, 1999



Supervisor's Certificate (Sample)

I, hereby, certify that Dr. _____ having Enrolment
Number : _____ & RTMC Registration
Number: _____ has been working under my direct
supervision with effect from :

(date) _____ to (date) _____

in the Department: _____ Unit: _____

of Training institution: _____

in the City of: _____

The enclosed Dissertation titled:

_____ was prepared according to the “FCPS Dissertation - Guidelines” under
my direct supervision. I have read the Dissertation and have found it
satisfactory for FCPS part II examination in the subject.

Signature of the Supervisor: _____

Name of the Supervisor: _____

Designation: _____ Date: _____

Official stamp:



Acknowledgement(s)

- May be included at the end of the dissertation or after the table of contents.
- May acknowledge
Technical help
(separate paragraph),
Financial and material support.



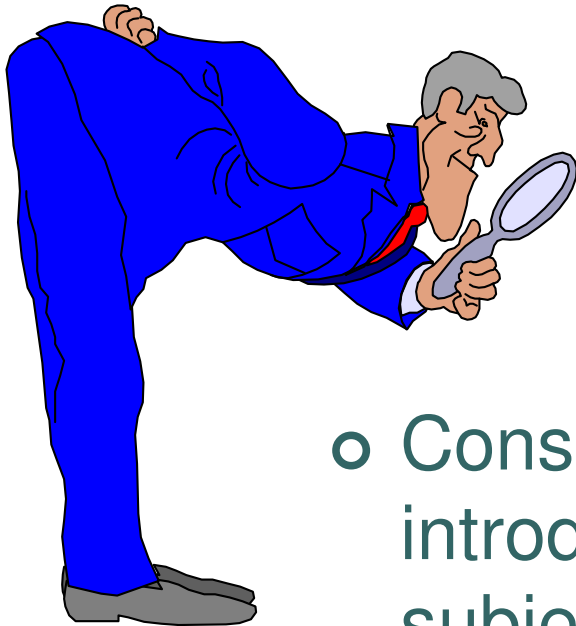


TABLE OF CONTENTS

- Heading
 - Subheading
 - Annexes
- Page #**

Page #s

- Roman numerals till start of summary
- Arabic numerals from summary till annexes



Abstract

- Consists of 250 words covering the introduction, objectives, design, setting, subjects & methods, results, & conclusion of the study
- Three to ten key words identified and written at the end



INTRODUCTION

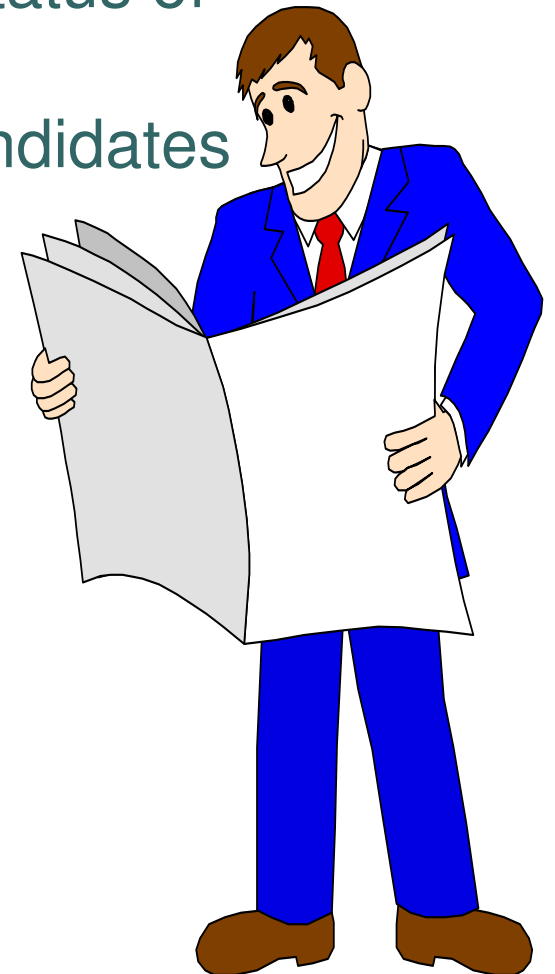
(Why did you study the problem?)

- Identify problems & issues (magnitude of the problem)
- Background (Why is it important to conduct the study?)
- What previous work has been undertaken on the topic?
- What new will your study be evaluating and its impact



Review Of The Literature

- Comprehensive review of the current status of knowledge on the selected topic.
- Collective review and critique in the candidates own words (not copied).
- References of the last 5 years (older, relevant and historical references can be used).
- Review of the local literature must be included.
- Literature cited in MedLine, ExtraMed & journals approved by PMDC.





Objective(s) of Study

- Should be stated in clear & quantifiable words
- Should be the same as stated in synopsis



Operational Definition

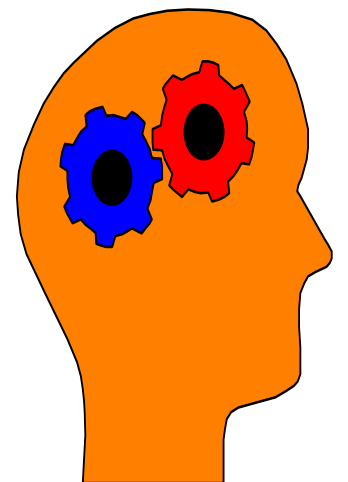
- Should clearly define and indicate means of measurement / determination of all exposure / outcome (dependent & independent) variables indicated in the objectives.



MATERIALS & METHODS

(What did you do? How did you do it?)

- Study designs
- Study population
- Procedures / apparatus used
- Methods of gathering & analyzing data
- Statistical analysis undertaken
(Write in past tense)





Results



- The result should be in logical sequence.
- Main result should be stated first.
- Summarization of the data collected and statistical analysis made should be given in this section.

Discussion



- Results of the study should be examined, interpreted.
- Implications and limitations described.



Conclusion(s)

- It describes the inferences drawn on the basis of the results of study.
- Should be linked with the objectives and purpose of the study.



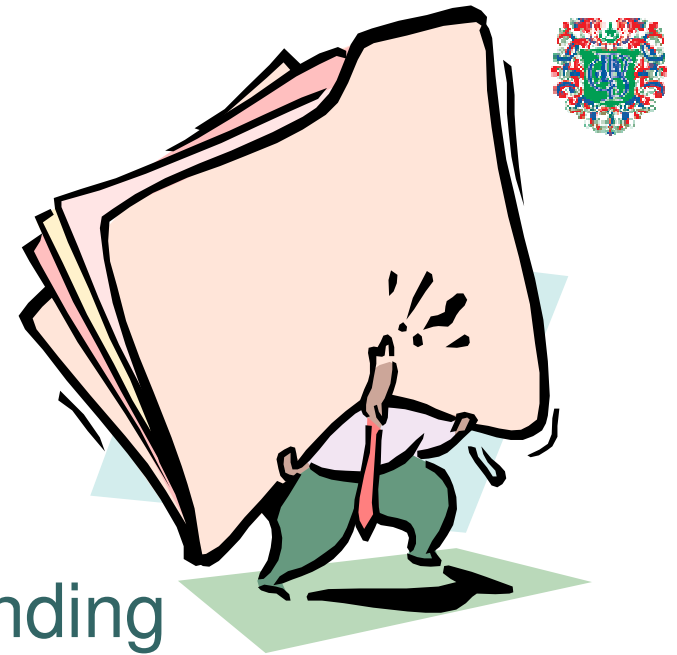


References

- Citation in the text
 - serially numbered
- Listing references
 - at the end in Vancouver style.

Annexes

- should be added if they increase the understanding or evaluation of the study.
- All annexes should be serially numbered and referred to at appropriate places in the body of dissertation.





REFERENCE WRITING



● ● ● | **Reference writing**

- Vancouver Style
- International Committee of Medical Journal Editors
Uniform Requirements for Manuscripts Submitted to
Biomedical Journals



Articles in Journals

All authors should be listed. If there are six or less

- Halpern SD, Ubel PA, Caplan AL. Solid-organ transplantation in HIV-infected patients. N Engl J Med. 2002 Jul 25;347(4):284-7.



Articles in Journals

If a journal carries continuous pagination throughout a volume the month and issue number may be omitted.

- Halpern SD, Ubel PA, Caplan AL. Solid-organ transplantation in HIV-infected patients. N Engl J Med. 2002;347:284-7.



Articles in Journals

More than six authors

- List the first six authors followed by et al
- Rose ME, Huerbin MB, Melick J, Marion DW, Palmer AM, Schiding JK, et al.
Regulation of interstitial excitatory amino acid concentrations after cortical contusion injury. Brain Res. 2002;935:40-6.



Organization as author

Diabetes Prevention Program Research Group. Hypertension, insulin, and proinsulin in participants with impaired glucose tolerance. Hypertension. 2002;40 :679-86.



No author given

- 21st century heart solution may have a sting in the tail. BMJ.2002;325:184.



Volume with supplement

- Geraud G, Spierings EL, Keywood C. Tolerability and safety of frovatriptan with short- and long-term use for treatment of migraine and in comparison with sumatriptan. Headache. 2002;42 Suppl 2:S93-9.



Issue with supplement

- Glauser TA. Integrating clinical trial data into clinical practice. *Neurology*. 2002; 58(12 Suppl 7):S6-12.



Volume with part

Abend SM, Kulish N. The psychoanalytic method from an epistemological viewpoint. Int J Psychoanal. 2002;83(Pt 2):491-5.



Issue with part

- Ahrar K, Madoff DC, Gupta S, Wallace MJ, Price RE, Wright KC. Development of a large animal model for lung tumors. J Vasc Interv Radiol. 2002;13(9 Pt 1):923-8.



Issue with no volume

- Banit DM, Kaufer H, Hartford JM. Intraoperative frozen section analysis in revision total joint arthroplasty. Clin Orthop. 2002;(401):230-8.



Book References

- Personal author(s)
- Murray PR, Rosenthal KS, Kobayashi GS, Pfaller MA. Medical microbiology. 4th ed. St. Louis: Mosby; 2002.



Editor(s), compiler(s) as author

- Gilstrap LC 3rd, Cunningham FG, VanDorsten JP, editors. Operative obstetrics. 2nd ed. New York: McGraw-Hill; 2002.



Author(s) and editor(s)

- Breedlove GK, Schorfheide AM. Adolescent pregnancy. 2nd ed. Wieczorek RR, editor. White Plains (NY): March of Dimes Education Services; 2001.



Chapter in a book

- Meltzer PS, Kallioniemi A, Trent JM. Chromosome alterations in human solid tumors. In: Vogelstein B, Kinzler KW, editors. The genetic basis of human cancer. New York: McGraw-Hill; 2002. p. 93-113.



Dissertation

- Borkowski MM. Infant sleep and feeding: a telephone survey of Hispanic Americans [dissertation]. Mount Pleasant : Central Michigan University; 2002.



Journal article on the Internet

- Abood S. Quality improvement initiative in nursing homes: the ANA acts in an advisory role. Am J Nurs [serial on the Internet]. 2002 Jun [cited 2002 Aug 12];102(6):[about 3 p.]. Available from: <http://www.nursingworld.org/AJN/2002/june/Wawatch.htm>