How to select study subjects using Sampling Technique

Objectives:

To understand:

- Why we use sampling methods
- Definitions of few concepts
- Sampling and non-sampling methods
- And able to use sampling methods appropriately

Team Members: Weam Babaier - Njoud Alenezy – Lama Alfawzan

Team Leaders: Rawan Alwadee & Mohammed ALYousef

Revised By: Basel almeflh

Doctor: Dr.Shafi Ahemed

Resources:

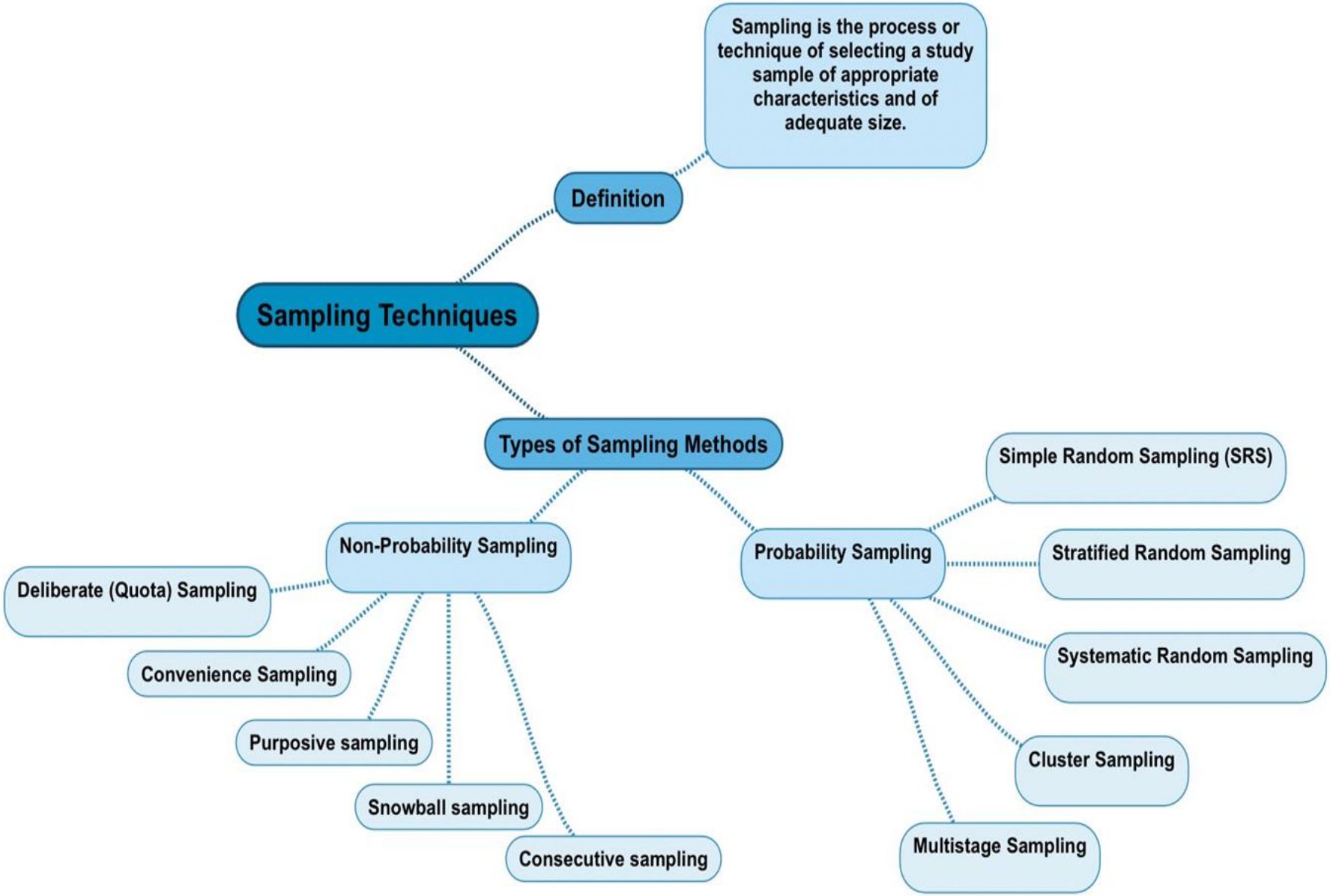
- 436 Lecture Slides + Notes

Important – Notes

436researchteam@gmail.com

Editing file

Feedback link


Sampling is the process or technique of selecting a study sample of appropriate characteristics and of adequate size.
Definition
Sampling Techniques
Types of Sampling Methods
Non-Probability Sampling
Deliberate (Quota) Sampling
Convenience Sampling
Purposive sampling
Snowball sampling
Consecutive sampling
Probability Sampling
Simple Random Sampling (SRS)
Stratified Random Sampling
Systematic Random Sampling
Cluster Sampling
Multistage Sampling

Sampling:

Sampling is the process or technique of selecting a study sample of appropriate characteristics and of adequate size.

Sampling in Epidemiology

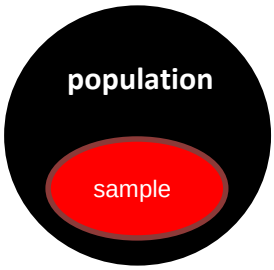
We studied the whole population only on one situation which is population census.it means trained investigator go door to door and collect information of the population country

• Why Sample?

- Unable to study all members of a population
- Reduce selection bias
- Save time and money
- Measurements may be better in sample than in entire population
- Feasibility

➢ **Definitions:**

❖ **Population** – group of things (people) having one or more common characteristics
-a set which includes all measurements of interest to the researcher
(The collection of all responses, measurements, or counts that are of interest)
Population is not only human it could be animals or blood samples.



❖ **Sample** – representative subgroup of the larger population

- Used to estimate something about a population (generalize)
- Must be similar to population on characteristic being investigated

-A subset of the population. you can take a sample from riyadh and generalize through KSA

❖ **Sampling Frame** ex. class: students list
This is the complete list of sampling units in the target population to be subjected to the sampling procedure.
Completeness and accuracy of this list is essential for the success of the study.

❖ **Sampling Units** ex. each student
These are the individual units / entities that make up the frame just as elements are entities that make up the population.

❖ **Sampling Error**
This arises out of random sampling and is the discrepancies between sample values and the population value.
discrepancies means incorrect or lack of compatibility

❖ **Sampling Variation**
Due to infinite variations among individuals and their surrounding conditions.
Produce differences among samples from the population and is due to chance

Example: In a clinical trial of 200 patients we find that the efficacy of a particular drug is 75% If we repeat the study using the same drug in another group of similar 200 patients we will not get the same efficacy of 75%. It could be 78% or 71%.
"Different results from different trials though all of them conducted under the same conditions" there will be variability

Representativeness (validity)

A sample should accurately reflect distribution of relevant variable in population

- Person e.g. age, sex
- Place e.g. urban vs. rural
- Time e.g. seasonality

- Representativeness essential to generalise & Ensure representativeness before starting & Confirm once completed.

adequate sample size plus technique would help in generalizability ex. DM in riyadh can be generalized to KSA
(precision also called reliability)

Illustration of the difference between precision and accuracy

Validity of a Study

Two components of validity:

➢ **Internal validity**

- A study is said to have internal validity when there have been proper selection of study group and a lack of error in measurement.(relate to (how you select samples what is sample and measurement))

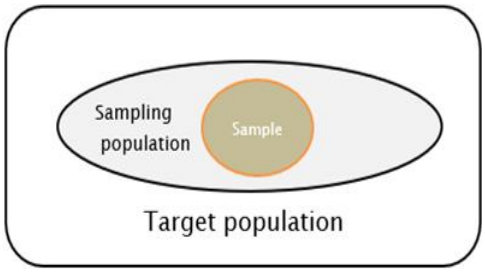
For example, it is concerned with the appropriate measurement of exposure, outcome, and association between exposure and disease.

➢ **External validity**

- External validity implies the ability to generalize beyond a set of observations to some universal statement.

here your sampling technique plays a big role and you can generalize. internal validity ---->method. external validity ---> generalizability

Sampling and representativeness



Target Population ➔ Sampling Population ➔ Sample

How to sample?

In general, 2 requirements

1. Sampling frame must be available, otherwise develop a sampling frame.
2. Choose an appropriate sampling method to draw a sample from the sampling frame.

The Sampling Design Process

probability sampling:

1.Simple Random Sampling: (each one has a chance of being selected- equal probability)

·Equal probability

·Techniques:

- Lottery method all names in a drum and choose NOT SCIENTIFIC
- Table of random numbers

·Advantage:

- Most representative group

·Disadvantage:

- Difficult to identify every member of a population

Table of random numbers

6 8 4 2 5 7 9 5 4 1 2 5 6 3 2 1 4 0

5 8 2 0 3 2 1 5 4 7 8 5 9 6 2 0 2 4

3 6 2 3 3 3 2 5 4 7 8 9 1 2 0 3 2 5

9 8 5 2 6 3 0 1 7 4 2 4 5 0 3 6 8 6

Random Number table:

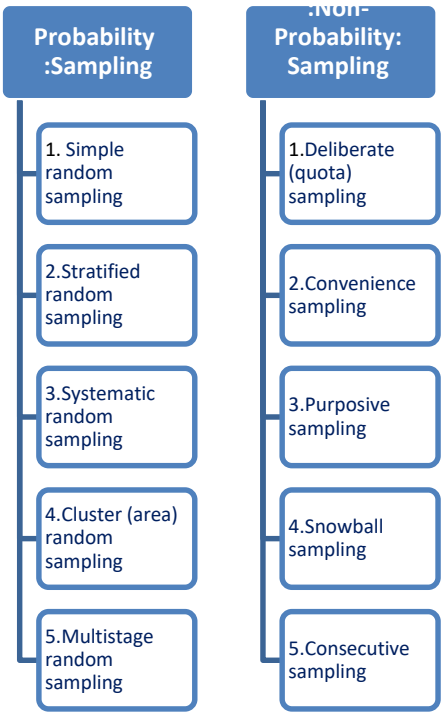
1	2	3	4	5
49486	93775	88744	80091	92732
94860	36746	04571	13150	65383
10169	95685	47585	53247	60900
12018	45351	15671	23026	55344
45611	71585	61487	87434	07498
89137	30984	18842	69619	53872
94541	12057	30771	19598	96069
89920	28843	87599	30181	26839
32472	32796	15255	39636	90819

How to select a simple random sample?

1. Define the population
2. Determine the desired sample size
3. List all members of the population or the potential subjects

✓ For example: 4th grade boys who have demonstrated problem behaviors

Sampling Methods:



Simple random sampling

➢ Estimate hemoglobin levels in patients with sickle cell anemia

1. Determine sample size
2. Obtain a list of all patients with sickle cell anemia in a hospital or clinic
3. Patient is the sampling unit
4. Use a table of random numbers to select units from the sampling frame
5. Measure hemoglobin in all patients
6. Estimate the levels (normal & abnormal) of hemoglobin

Let's select 10 boys from the list

Potential Subject Pool

1. Ahamed
2. Munir
3. Khalid
4. Ameer
5. Junaid
6. Khadeer
7. Shaffi
8. Rafi
9. Ghayas
10. Fayaz

11. Riyaz
12. Yaseen
13. Jaffar
14. Sattar
15. Ghouse
16. Imran
17. Khaleel
18. Shabu
19. Shanu
20. Javid

21. Fahad
22. Iqbal
23. Jabbar
24. Aziz
25. Anwar
26. Shohail
27. Shohaib
28. Rehaman
29. Naeem
30. Rahim

So our selected subjects are with numbers 10, 22, 24, 15, 6, 1, 25, 11, 13, & 16.

1. Ahamed
2. Munir
3. Khalid
4. Ameer
5. Junaid
6. Khadeer
7. Shaffi
8. Rafi
9. Ghayas
10. Fayaz

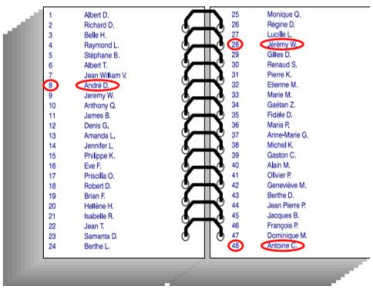
11. Riyaz
12. Yaseen
13. Jaffar
14. Sattar
15. Ghouse
16. Imran
17. Khaleel
18. Shabu
19. Shanu
20. Javid

21. Fahad
22. Iqbal
23. Jabbar
24. Aziz
25. Anwar
26. Shohail
27. Shohaib
28. Rehaman
29. Naeem
30. Rahim

2. Systematic random Sampling first number should be random,if your first number is not random it called systematic non random size

- Technique
  - Use "system" to select sample (e.g., every 5th item in alphabetized list, every 10th name in phone book)
- Advantage
  - Quick, efficient, saves time and energy
- Disadvantage
  - Not entirely bias free; each item does not have equal chance to be selected
  - System for selecting subjects may introduce systematic error
  - Cannot generalize beyond population actually sampled

Example



➢ If a systematic sample of 500 students were to be carried out in a university with an enrolled population of 10,000, the sampling interval would be:

➢ I = N/n = 10,000/500 =20

➢ All students would be assigned sequential numbers. The starting point would be chosen by selecting a random number between 1 and 20. If this number was 9, then the 9th student on the list of students would be selected along with every following 20th student. The sample of students would be those corresponding to student numbers 9, 29, 49, 69,..... 9929, 9949, 9969 and 9989.

*select a random starting point and then select every K (th) subject in the population

3. Stratified Random Sampling

➢ Technique
- Divide population into various strata all demographic variables are stratification so we stratify to remove confounding bias
- Randomly sample within each strata
- Sample from each strata should be proportional

➢ Advantage
- Better in achieving representativeness on control variable

➢ Disadvantage
- Difficult to pick appropriate strata
- Difficult to Identify every member in population

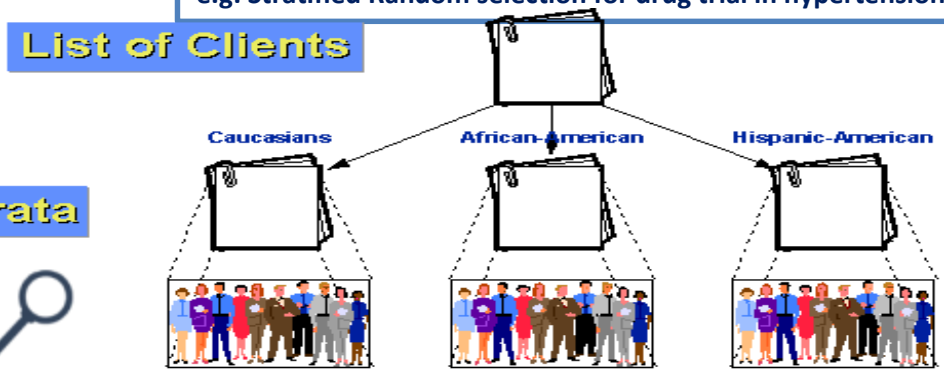
✓ Divide the population into at least two different groups with common characteristics, then draw subjects randomly from each group (group is called strata or stratum)

Stratified random sample

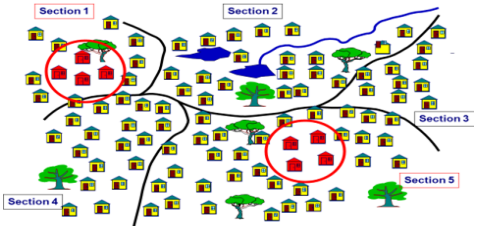
- Assess dietary intake in adolescents

1. Define three age groups: 11-13, 14-16, 17-19
2. Stratify age groups by sex
3. Obtain list of children in this age range from schools
4. Randomly select children from each of the 6 strata until sample size is obtained
5. Measure dietary intake

e.g. Stratified Random selection for drug trial in hypertension



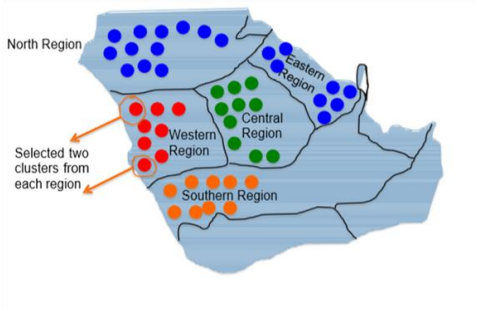
4. Cluster (Area) random sampling



- Randomly select groups (cluster) – all members of groups are subjects
- Appropriate when
  - you can't obtain a list of the members of the population
  - have little knowledge of population characteristics
  - Population is scattered over large geographic area

  simple---> random sampling

  if large ones ----> cluster sampling



- Advantage
  - More practical, less costly
- Conclusions should be stated in terms of cluster (sample unit – school)
- Sample size is number of clusters

5. Multistage random sampling

- Stage 1

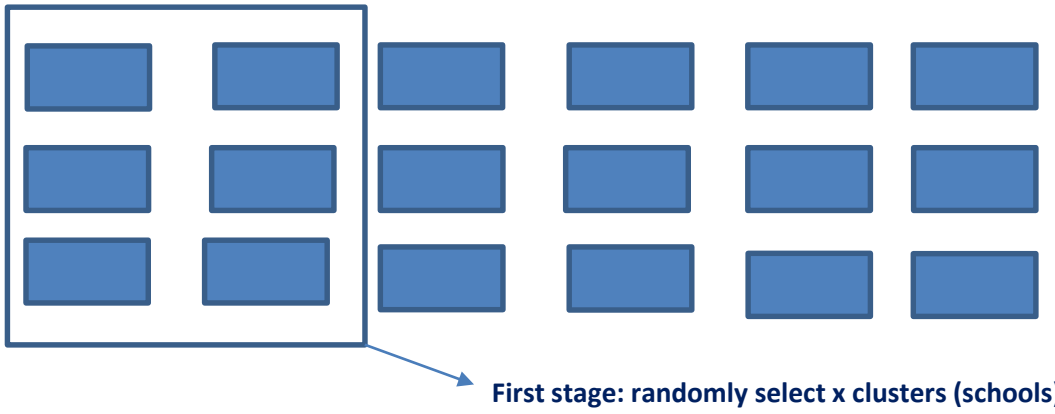
  --randomly sample clusters (schools)
- Stage 2

  --randomly sample classrooms from the schools selected
- Stage 3

  --random sample of students from class rooms

  It maybe 4 5 depends on the situation.

Do not get confused between cluster and stratified variables (all person characteristics of study subject).cluster could be regions of Riyadh (north south west) school AREAS.
Multistage of random sampling is extension of cluster random sampling multistage go to cluster and do again random sampling

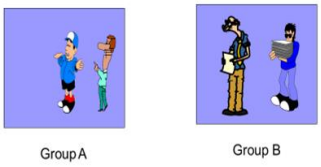


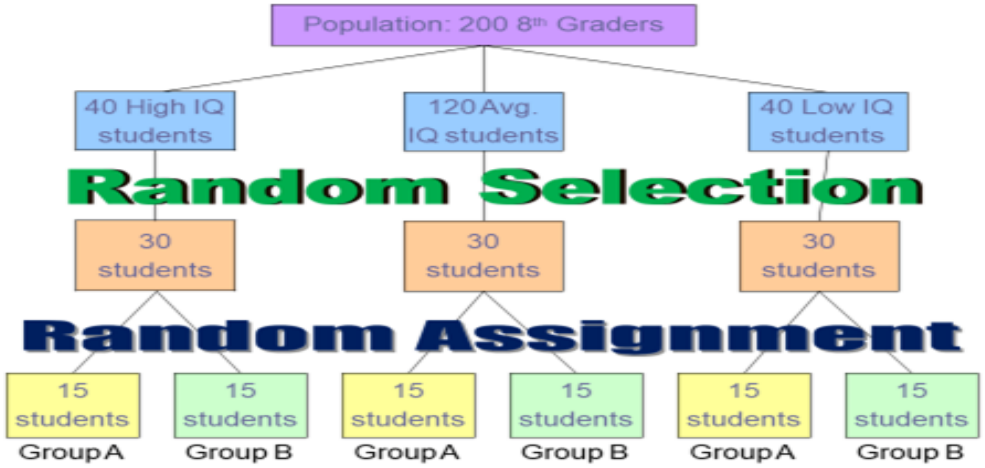
First stage: randomly select x clusters (schools)

Second stage: within each school, randomly select y clusters (class rooms)

Third stage: randomly select x number of people from the classroom

Random . . .

Random Assignment	Random Selection
• ¨Random Assignment = every member of the sample (however chosen) has an equal chance of being placed in the experimental group or the control group. • Random assignment allows for individual differences among test participants to be averaged out.	Random Selection = every member of the population has an equal chance of being selected for the sample
Deciding which group or condition each subject will be part of 	Choosing which potential subjects will actually participate in the study 



Non- probability sampling:

1. Deliberate (Quota) Sampling

- Similar to stratified random sampling
- Technique
  - Quotas set using some characteristic of the population thought to be relevant
  - Subjects selected non-randomly to meet quotas (usu. convenience sampling)
- Disadvantage
  - selection bias
  - Cannot set quotas for all characteristics important to study

2. Convenience Sampling

- "Take them where you find them" - nonrandom
- Intact classes, volunteers, survey respondents (low return), a typical group, a typical person
- Disadvantage: Selection bias Most disadvantage

3. Purposive Sampling

- Purposive sampling (criterion-based sampling)
  - Establish criteria necessary for being included in study and find sample to meet criteria.
- Solution: Screening
  - Obtain a sample of a larger population and then those subjects that are not members of the desired population are screened or filtered out. no methods of convenience

EX: want to study smokers but can't identify all smokers

another ex. if i want to measure med student BMI plus normal stat, obese go of class

4. Snowball Sampling

In snowball sampling, an initial group of respondents is selected.

➢ After being interviewed, these respondents are asked to identify others who belong to the target population of interest.

➢ Subsequent respondents are selected based on the referrals. no random methods

5. Consecutive sampling

➢ Outcome of 1000 consecutive patients presenting to the emergency room with chest pain

➢ Natural history of all 125 patients with HIV-associated TB during 5 year period

Explicit efforts must be made to identify and recruit ALL persons with the condition of interest

Choosing probability vs. non-probability sampling method

Probability sampling	Evaluation Criteria	Non-probability sampling
Conclusive	*Nature of research*	Exploratory
Larger sampling errors	*Relative magnitude sampling vs. non-sampling error*	Larger non-sampling error
High [Heterogeneous]	*Population variability*	Low [Homogeneous]
Favorable	*Statistical Considerations*	Unfavorable
High	*Sophistication Needed*	Low
Relatively Longer	*Time*	Relatively shorter
High	*Budget Needed*	Low

Homogenous ,Why? Because our sample is not random.

In Conclusion,

For any research, based on its study design and objectives an appropriate random sampling technique should be used.

THE END