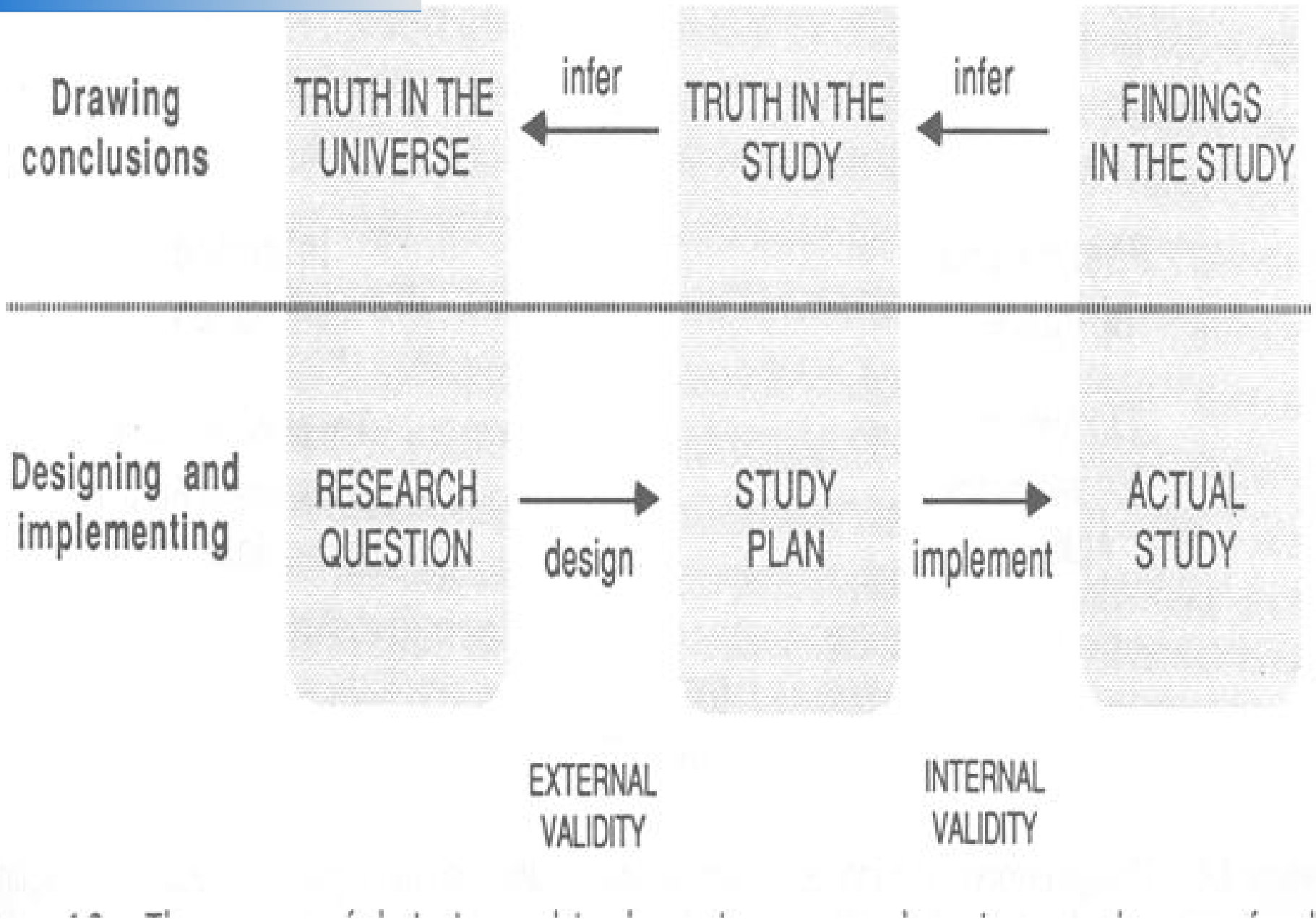
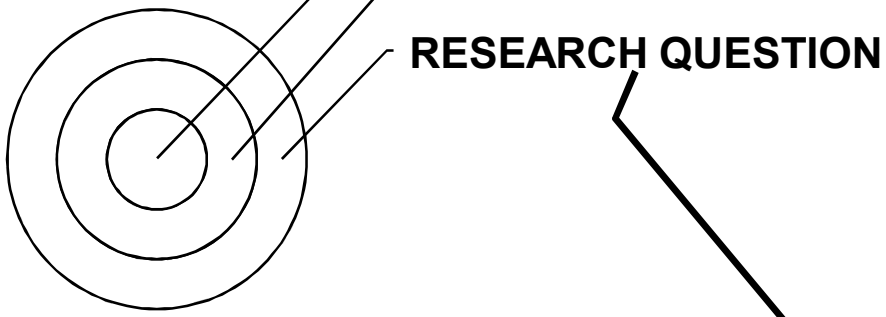
# Research Question, Hypothesis, Variables

**Dr. R.M. Pandey**
Prof & Head
Department of Biostatistics
A.I.I.M.S., New Delhi
rmpandey@yahoo.com

# gning & Implementing a Research Project





HYPOTHESIS

STUDY QUESTION

RESEARCH QUESTION

The knowledge of research methodology increases after two days of sessions when evaluated through scores on a pretest and posttest

Whether two days of sessions Increase immediate understand -ing of research methodology.

Whether a research training session increases participants knowledge on research methodology.

# The research question

- All studies should start with a *research question* that addresses what the investigator would like to know

- Goal is to find an important research question that can be developed into a feasible and valid study plan

**Research Question**

↓

**Study Design**

↓

**Data Collection**

↓

**Data Analysis**

↓

**Interpretation**

# Sections of the study protocol

- *Research question*
- **Significance (background)**
- **Study design**
- **Study population and sampling**
- **Variables and measurements**
- **Statistical issues**
- **Ethical issues**
- **Quality control and data management**

# Origins of a research question

- Mastering the literature
  - ï Scholarship -- be a scholar and know the literature
  - ï Get a mentor
- Be alert to new ideas and techniques
  - ï Importance of teaching
- Be imaginative

# Narrowing and Clarifying

- Narrowing, clarifying, and even redefining your questions is essential to the research process.

- Forming the right ôquestionsô should be seen as an iterative process that is informed by reading and doing at all stages.

# The research question

- Format for descriptive study

  ñIn a population of [*study population*], what is the prevalence (or mean, median, etc.) of [*outcome variable*]?ò

- Format for analytic study

  ñIn a population of [study population], is [predictor variable] associated with [outcome variable]?ò

# FINER characteristics

- *F*easible
- *I*nteresting to the investigator
- *N*ovel
- *E*thical
- *R*elevant

# racteristics criteria for feasibility

- Adequate number of subjects
- Adequate technical expertise
- Affordable in time and money
- Manageable in scope

# acteristics criteria for Interesting

- Should be of academic interest

# racteristics Criteria for Novelty

- Confirms or refutes previous findings
- Extends previous findings
- Provides new findings

# racteristics criteria for Relevance

- To scientific knowledge
- To clinic, public health or health policy
- To future research directions

# Problems and solutions: Research plan is not *FINER*

- Not feasible
  - ï Too broad
  - ï Not enough subjects available
  - ï Methods beyond skill of investigator
  - ï Too expensive
- Not interesting, novel or relevant
- Uncertain ethical suitability

# Problems and solutions: Research plan is not feasible

- Too broad
  - Smaller set of variables
  - Narrow the question
- Not enough subjects available
  - Expand inclusion criteria
  - Modify exclusion criteria
  - Add other sources of subjects
  - Lengthen the time frame for entry into study
  - Use strategies to decrease sample size

# oblems and solutions: Research plan is not feasible

- Methods beyond skill of investigator
  - ï Collaborate with colleagues who have skills
  - ï Consult experts and review literature for alternative methods
  - ï Learn the needed skills yourself
- Too expensive
  - ï Consider less costly study design
    - Fewer subjects and measurements
    - Less extensive measurement
    - Fewer follow-up visits

# Problems and solutions: Research plan is not *FINER*

- Not interesting, novel or relevant
  - ï Consult with mentor
  - ï Modify the research question
- Uncertain ethical suitability
  - ï Consult with institutional review board (ethics committee)
  - ï Modify the research question

# Problems and solutions

- Study plan is vague
  - ï Write the research plan at an early stage
  - ï Get specific in the 1 to 2-page study plan
    - How the subjects will be sampled
    - How the variables will be measured

# Primary and Secondary Questions

- More than one primary question has advantages and disadvantages
- A sensible strategy is to establish single primary research question around which build the study plan
- This can be supplemented with secondary questions

# Five Important Terms

- Outcome variable
- Exposures
- Bias
- Confounder(s)
- Chance factor

# search Questions
## Primary/ Secondary

- What is the prevalence of a condition or mean of a characteristics?
- Correlation/Agreement between parameters
- What are the diagnostic characteristics of a candidate test (quantitative/categorical) with reference to a ñGold Standardò?.

- What  is incidence of an outcome?.
- What  are the predictors of an outcome?

- What are the risk factors associated with an outcome?

- Evaluation of a candidate intervention against a comparative (standard/control) intervention?.

# *pothesis:* Definition

**the word hypothesis is derived form the Greek words**

- **“hypo” means under**
- **“tithemi” means place**

**Under known facts of the problem to explain relationship between these**

- **.. is a statement subject to verification**
- **.. a guess but experienced guess based on some facts**
- **.. is a hunch, assumption, suspicion, assertion or an idea about a phenomena, relationship, or situation, the reality of truth of which one do not know**

**A researcher calls these assumptions, assertions, statements, or hunches and they become the basis of an inquiry.**

**In most cases, the hypothesis will be based upon either previous studies or the researcher’s own or someone else’s observations**

**Hypothesis is a conjectural statement of relationship between two or more variable**

# Definition

Hypothesis is proposition, condition or principle which is assumed, perhaps without belief, in order to draw its logical consequences and by this method to test its accord with facts which are known or may be determined (Webster's New International Dictionary of English).

A tentative statement about something, the validity of which is usually unknown

Hypothesis is proposition that is stated is a testable form and that predicts a particular relationship between two or more variable. In other words, id we think that a relationship exists, we first state it is hypothesis and then test hypothesis in the field

# Definition

A hypothesis is written in such a way that it can be proven or disproven by valid and reliable data ï  in order to obtain these data that we perform our study  )

A hypothesis may be defined as a tentative theory or supposition set up and adopted provisionally as a basis of explaining certain facts or relationships and as a guide in the further  investigation of other facts or relationships

# Characteristics

Hypotheses has the following characteristics:

- a tentative proposition
- unknown validity
- specifies relation between two or more variables

# Functions

**Bringing clarity to the research problem**

**Serves the following functions**

- ✓ provides a study with focus
- ✓ signifies what specific aspects of a research problem is to investigate
- ✓ what data to be collected and what not to be collected
- ✓ enhancement of objectivity of the study
- ✓ formulate the theory
- ✓ enable to conclude with what is true or what is false

# Characteristics

Simple, specific, and contextually clear

Capable of verification

Related to the existing body of knowledge

Operationalisable

# Typologies

Three types

- ✓ working hypothesis
- ✓ Null hypothesis
- ✓ Alternate hypothesis

# Example

- Working hypothesis: Population influences the number of bank branches in a town

- Null hypothesis (Ho): Population do not have any influence on the number of bank branches in a town.

- Alternate hypothesis ($H_1$): Population has significant effect on the number of bank branches in a town. A researcher formulates this hypothesis only after rejecting the null hypothesis.

# Typologies

**Null hypothesis**

A null hypothesis is formulated against the working hypothesis; opposes the statement of the working hypothesis

....it is contrary to the positive statement made in the working hypothesis; formulated to disprove the contrary of a working hypothesis

When a researcher rejects a null hypothesis, he/she actually proves a working hypothesis

In statistics, to mean a null hypothesis usually Ho is used.

For example,

Ho: There is no association between smoking and HTN

# Typologies

**Alternate hypothesis**

An alternate hypothesis is formulated when a researcher totally rejects null hypothesis

He/she develops such a hypothesis with adequate reasons

The notion used to mean alternate hypothesis is $H_1$ →Q>O

i.e., Q is greater than O

# Variables

- Definition: Variables are properties or characteristics of people or things that vary in quality or magnitude from person to person or object to object (Miller & Nicholson, 1976)
  - **Demographic characteristics**
  - **Anthropometric Characteristics**
  - **Nutritional variables**
  - **Clinical Variables**
  - **Behavioral variables**
  - **Abstract variables**
  - **Construct variables; etc.,**

- in order to be a variable, ***a variable must vary*** (e.g., not be a constant), that is, it must take on different values, levels, intensities, or states

xists because in health sciences things vary.

- Variables can be classified differently according to their utility and the vision of the scientific discipline.
- In medical research, an understanding of variables is important from both epidemiological (relational) and biostatistical perspectives.
- In epidemiology, variables are considered in terms of description or presentation, and relationships.
- From analysis perspective, allocation of specific variables into the specific classes is relative. It depends upon the level of knowledge, study questions and methodology.

# Outcome Variable

- Outcome; Dependent; Effect, Response

- In which investigator is actually interested

- | *Study* | *Exposure* | *Outcome* |
  |---|---|---|
  | 1 | SMK | HTN |
  | 2. | HTN | CVD |
  | 3. | CVD | Death |

# Exposure Variables

- Exposure; independent, covariate, factor
- Variable that is ***manipulated*** either by the researcher or by nature or circumstance
- These variables are also called ñstimulusò ñinputò or ñpredictoròvariables
- analogous to the ñcauseò in a cause-effect relationship

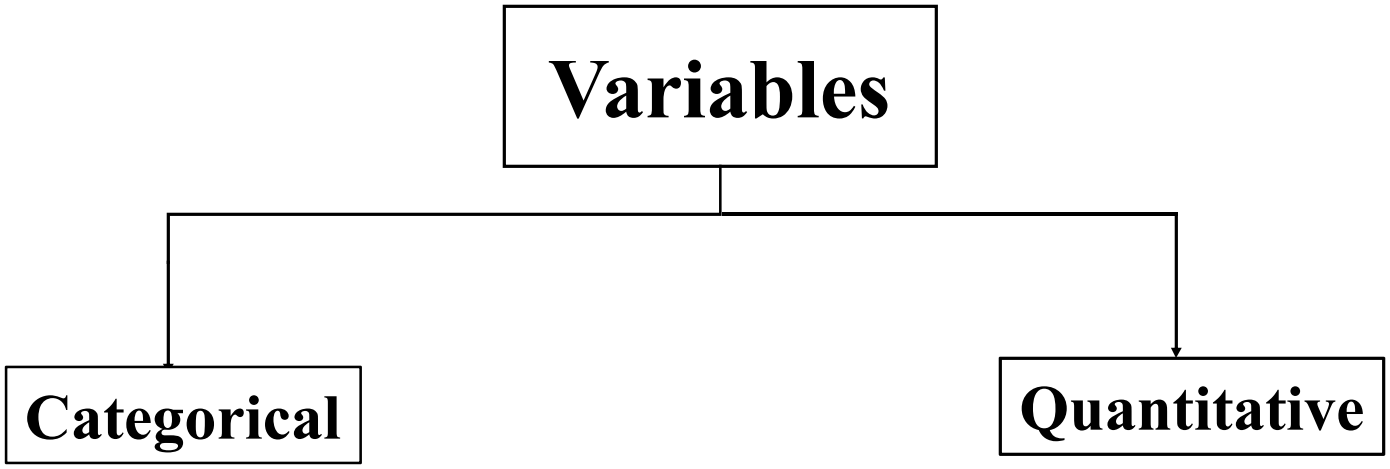
# nal studies: Not assigned

Modifiable/Non-Modifiable

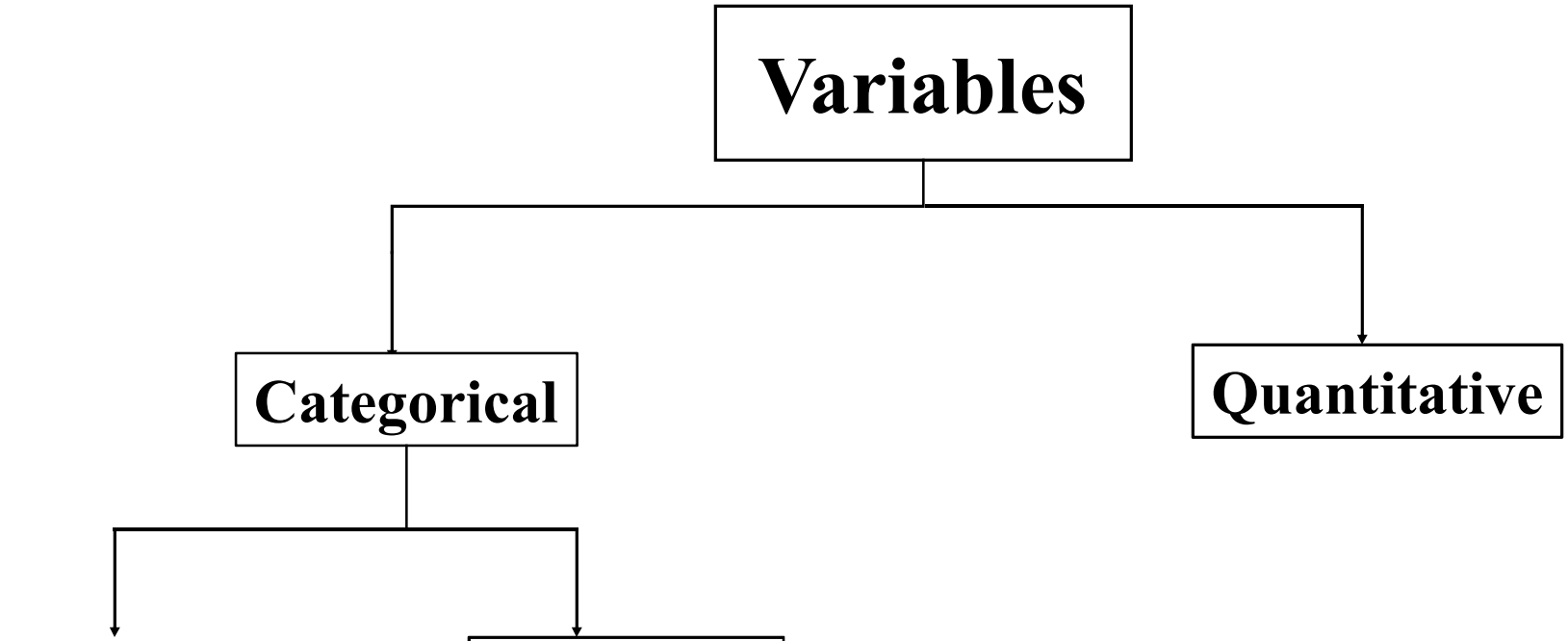
- Intervention studies : Assigned by the Investigator

- Any factor which has potential to influence the outcome would be an exposure; however, interested in one exposure at a time, other factors would be termed as other factors (Confounders; Effect modifier; intervening)

# pes of Variables: Level of Measurement



# pes of Variables: Level of Measurement




Gender

Smk

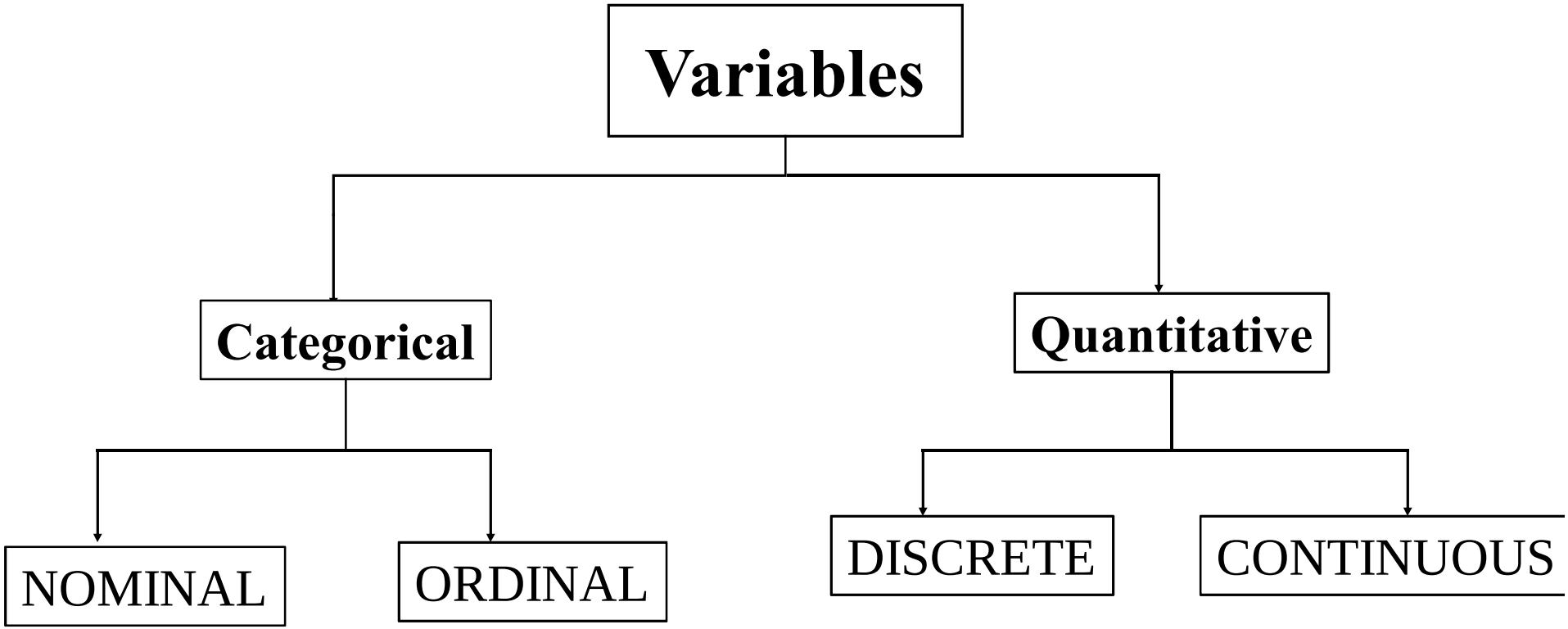
Treat groups

Cure

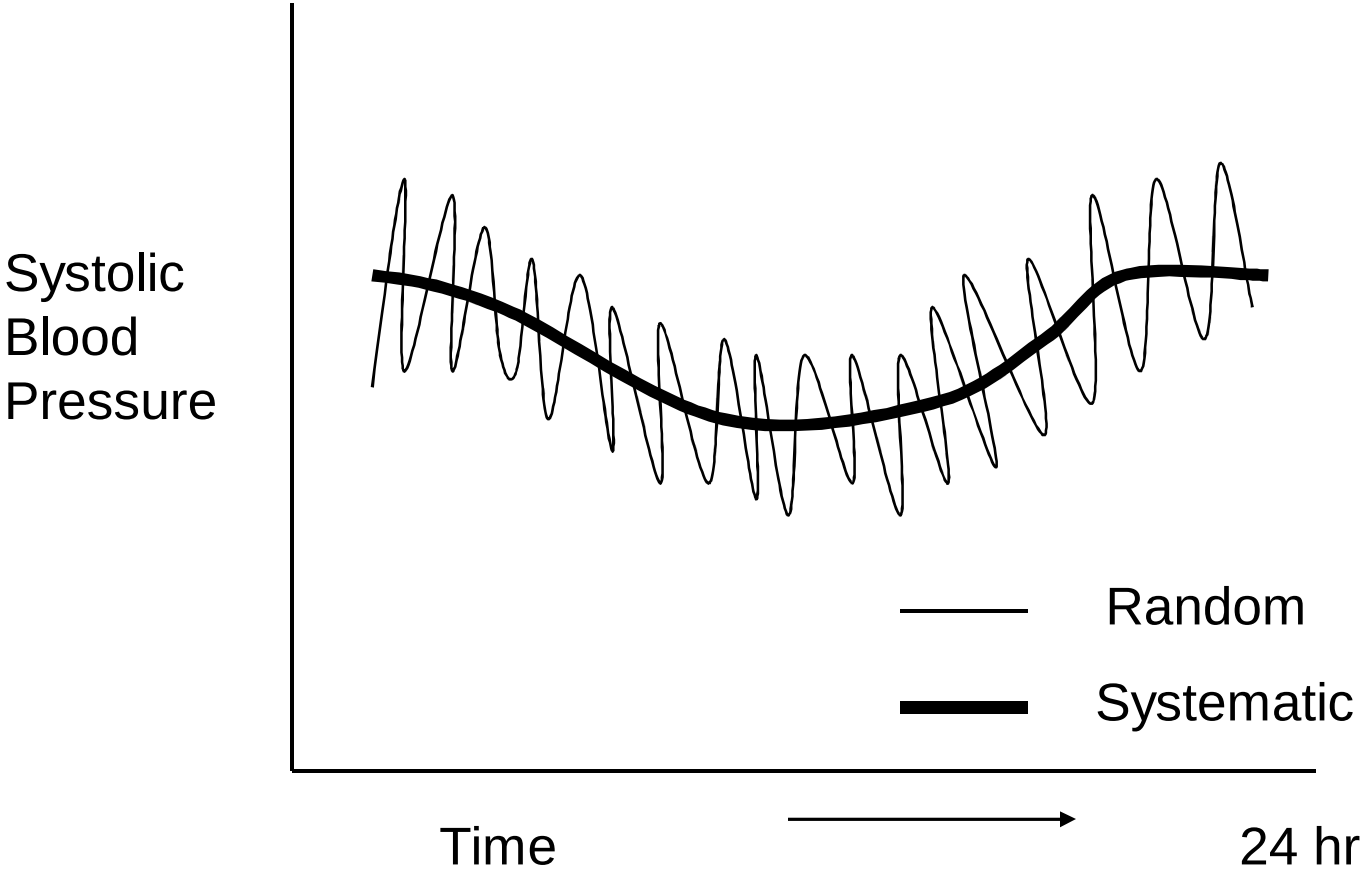
Severity of Disease

Dose groups

# pes of Variables: Level of Measurement



# Types of variability



# posure-Outcome Association : Possibilities

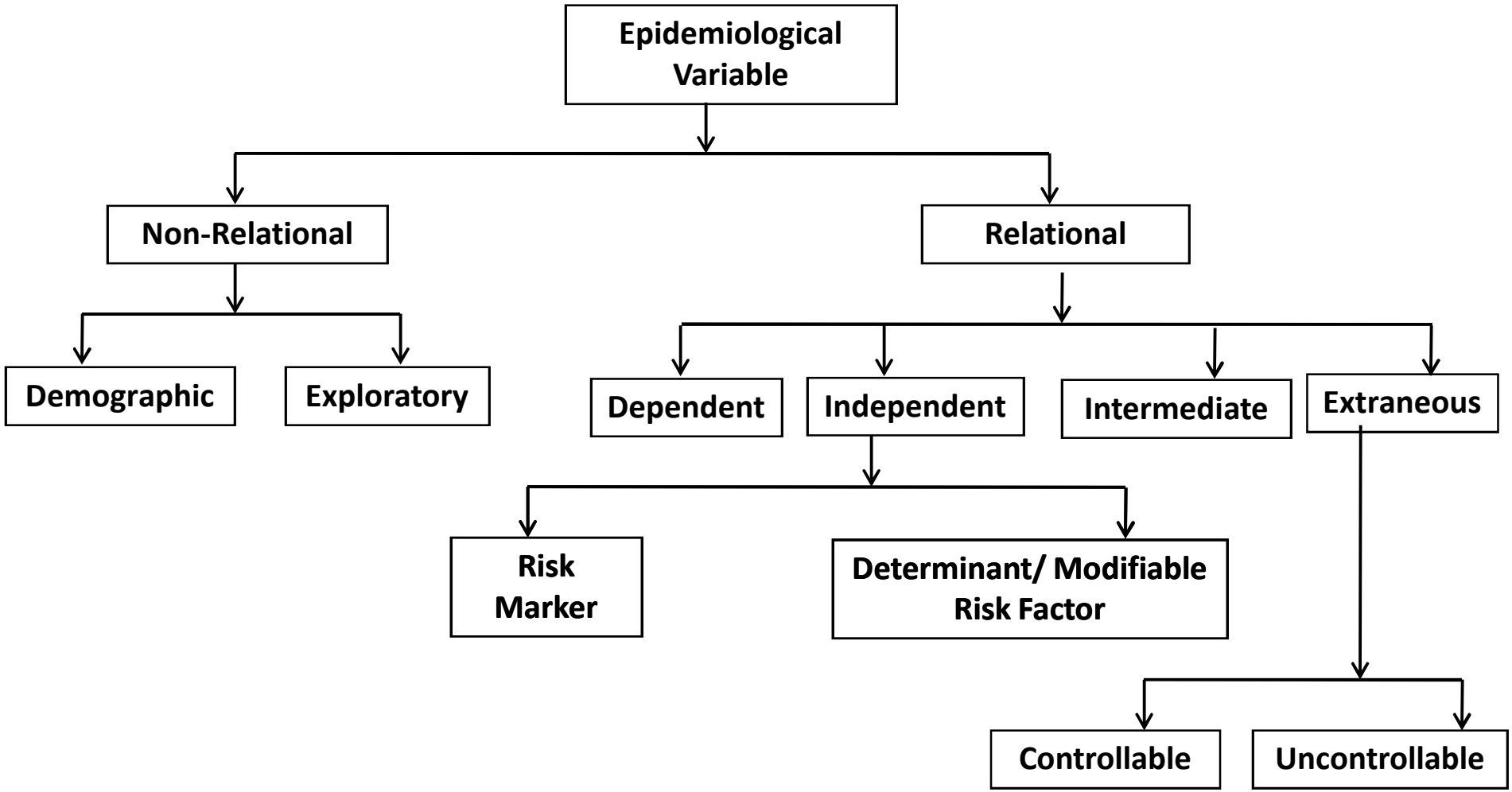
- The observed statistical association between a certain outcome and the hypothesized exposure could be the result of systematic errors in collection of data (sampling, disease and exposure ascertainment) or its interpretation
  - role of bias
- Or it could be due to the effect of additional variables that might be responsible for the observed association
  - role of confounding
- Or it could be just a matter of chance
- Or it could be a real association

# es of Variables: From Analysis Point of View

- Outcome
- Exposure
- Other factors
  - Confounder
  - Effect Modification
  - Intermediate

**From analysis perspective, allocation of specific variables into the specific classes is relative. It depends upon the level of knowledge, study questions and methodology.**

# cation of Variables in Epidemiology



# Questions??