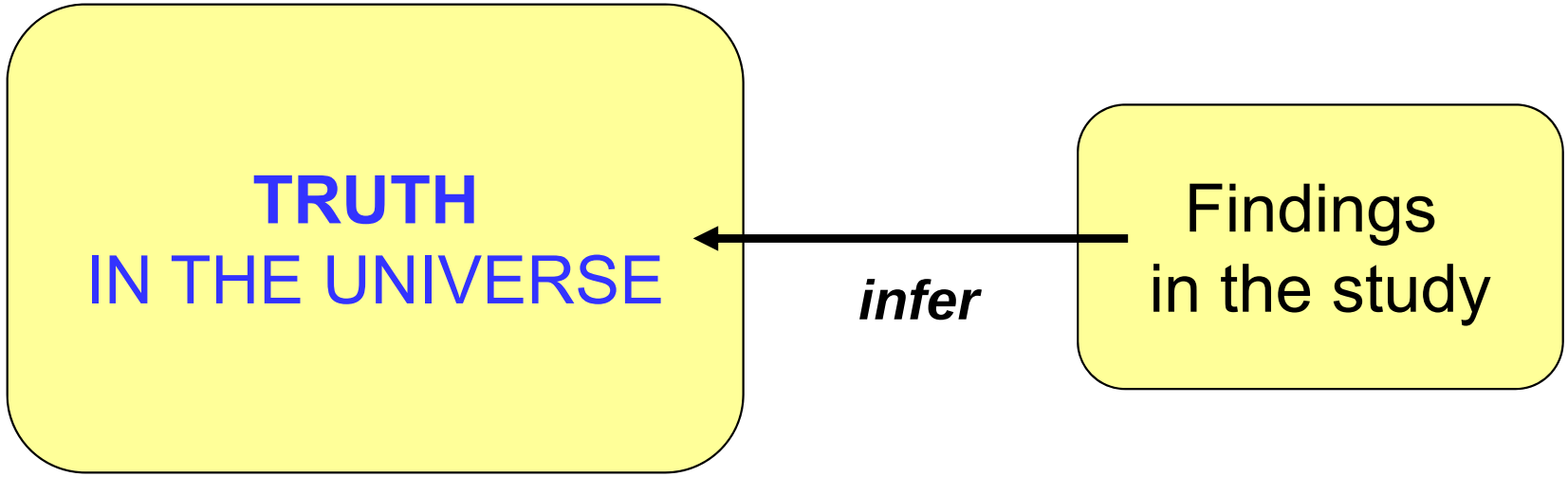
Learning Epidemiology
Basic Concepts in Epidemiology and Clinical Research

Trần Thế Trung, MD, MSc
Department of Endocrinology
University of Medicine and Pharmacy, Ho Chi Minh City

Contents

- The purpose of research
- Types of research design
  - Prospective vs. Retrospective
  - Cross-sectional vs. longitudinal study
- Research question
- Variables: barebones of quality research
  - Accuracy vs. Precision

The Physiology of Research



Research: Two viewpoints

(1) Anatomy of research – what it's made of

(the research question, design, subjects, measurements, sample size calculation…)

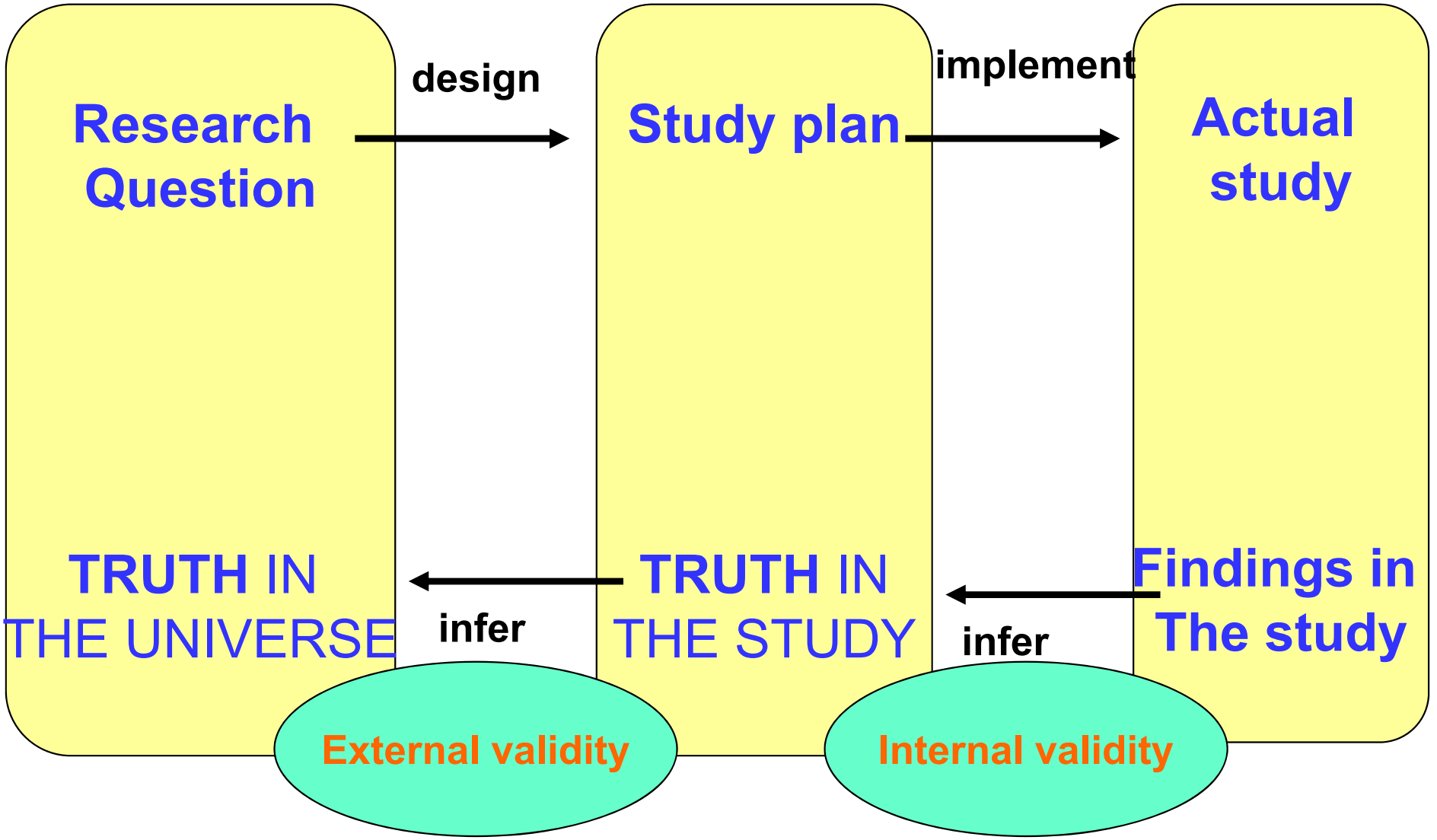
(2) Physiology of research – how it work.

- Internal validity: the events that happen in the study sample.

- External validity: generalizing these events to people outside the study.

- Threats: errors, random error and systematic error (= bias)

The Physiology of Research



Anatomy of research

The Research Question

- The objective of the study
- The uncertainty to resolve
- Must be *narrowed* (specific)
- Significance:
  - What is known at hand?
  - Why is the research question important?
  - What kind of answers will the study provide?

Anatomy of research

The Design

A complex issue

- Observational study >><< Interventional study
- Observational designs:
  - Case report
  - Case-series report
  - Cross-sectional study
  - Diagnostic test. Screening test.
  - Case-control study
  - Cohort study
- Interventional design

Anatomy of research

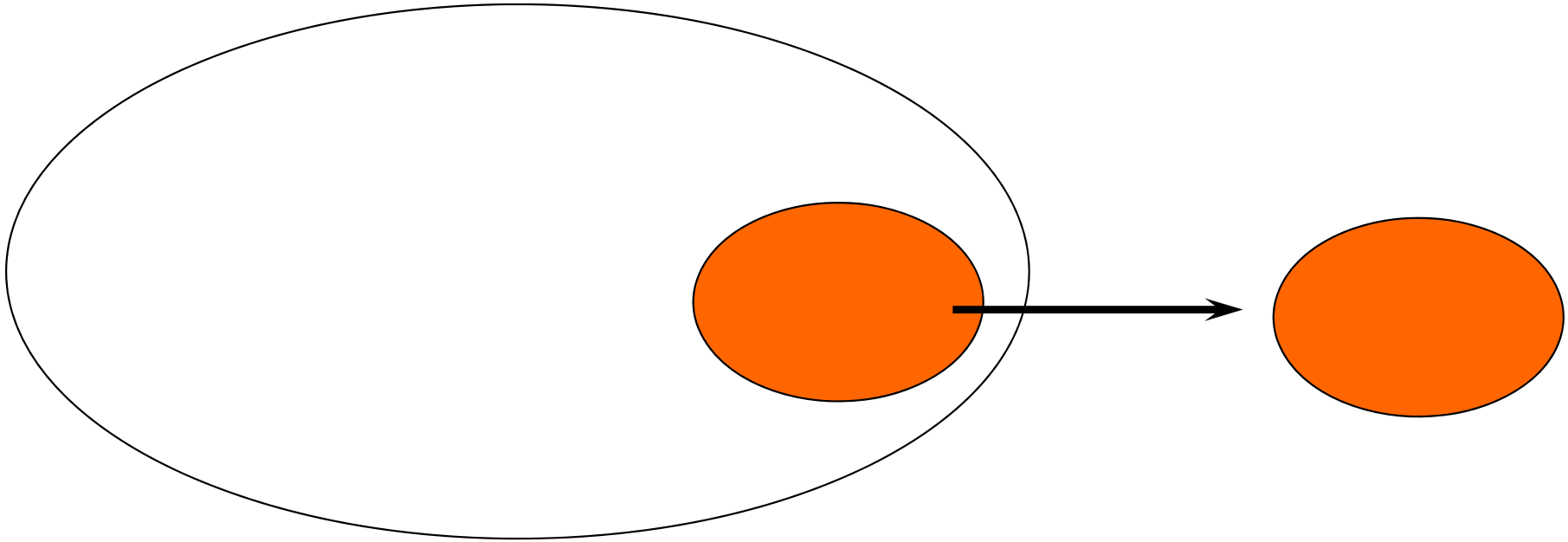
The Design

- Descriptive study >< Analytic study
- Cross-sectional >< Longitudinal
- Retrospective / Prospective / Historical Prospective / Retro-prospective
- Random / Nonrandom
- Blind / Unblind

Anatomy of research

Study Subjects

- Who ? => The target population (specific)
- How to recruit ? => Sampling



Anatomy of research

Variables

- Variables = Barebones of quality research
- Which variables are needed?
  - Predictor variables
  - Outcome variables

Anatomy of research

Statistical issue

- Hypothesis :
  - Some study (descriptive study) do not require a hypothesis.
  - Analytic studies and experiments : at least one main hypothesis.
- Sample size estimation:
  - Proportion
  - Difference between two means …

The Physiology of Research

- Designing the study
- Implementing the study
- Drawing causal inference

The Errors of Research

- No study is free of errors.
- The inferences are never perfectly valid.
- GOAL: maximize internal & external validity
- Errors:
  - Design phase
  - Implementation phase
  - Analysis phase

The Errors of Research

- Two type of Errors:
  - Random error = due to chance.
  - Systematic error = bias.

Research Question

Origins of a research question

- Build on experience (his own prior studies, his own works, … in the field).
- Mastering the published literature in an area of study.
- Senior scientist.

Origins of a research question

- Be alert to new ideas
- A skeptical attitude about prevailing beliefs
- New technologies
- Careful observation of patients

Origins of a research question

- Keep the imagination roaming
- Creativity
- Inspirations:
  - Colleague conversation
  - Brainstorming session
  - Preparing a lecture
  - Sitting and thinking
- Tenacity, until the problem have a resolution that feels comfortable.

A good research question: FINER

- Feasible:
  - Subjects (adequate number of subjects).
  - Technical expertise (adequate).
  - Cost in time and money (affordable).
  - Scope (manageable, narrow).
- Interesting (to the investigator)
- Novel (confirms, extends, provides new findings)
- Ethical
- Relevant (knowledge, policy, future research, …)

Developing the research question and study plan

- Write down the research question.
- Write down a brief outline of the study plan.
  - How the subjects will be sampled
  - How the variables will be measured

⇒ Problems (not FINER) and solutions.

⇒ Iterative process.

Number of question in a study

- Primary question(s)
- Secondary questions

=> A single primary question is favorable.

Estimate Sample Size: Hypotheses & Underlying Principles

Sample size

- How many subjects to sample?
- Sample size: too small, may fail to answer the **research question**.
- Too large: more difficult and costly.

=> *Appropriate* number.

- Estimate based on data (often guesses)
- Feasible ? Variables ? Any change?

=> Sample size should be estimated early!

Hypotheses

- **Research question** → research hypotheses --→ statistical tests.
- Elements of the study involved:
  - The sample
  - The design
  - Predictor variables
  - Outcome variables

Hypotheses

- Does any study need a hypotheses?

  Some do not need a hypotheses…
- Some have more than one hypotheses.

  Ex.
  - Prevalence of DM in CR workers.
  - Comparing the effectiveness of PTU and of Methimazole in treating Grave's disease.

Characteristics of a Good Hypotheses

- Simple (vs. complex)
  - Hyperglycemia -> Nephropathy
  - Hyperglycemia, HT, tobacco -> atherosclerosis, stroke and MI.
- Specific (vs. vague)
  - Exercise -> to lower cholesterol.
- In advance (vs. after-the-fact)

Statistical Principles

- Type I & Type II errors

Truth / Result	Null Hypo is true	Null Hypo is wrong
Accept Null	Correct	Type II
Different	Type I	Correct

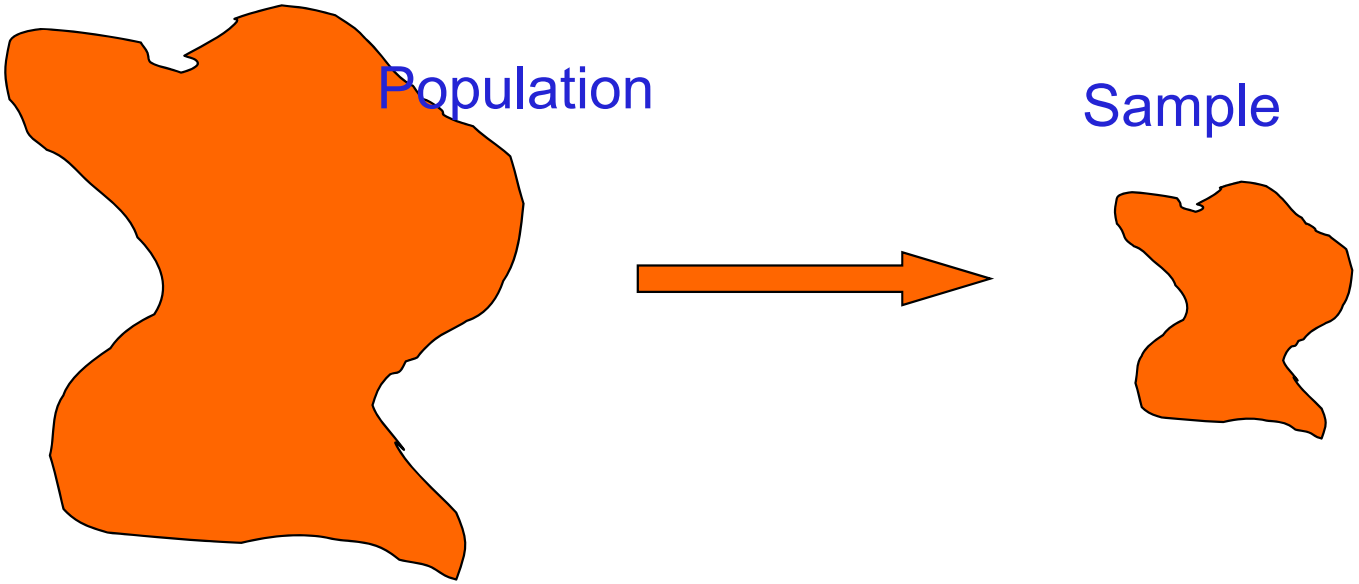
Sampling Technique

Random sampling

- Simple random sampling
- Cluster (random) sampling
- Stratified (random) sampling
- Systematic (random) sampling

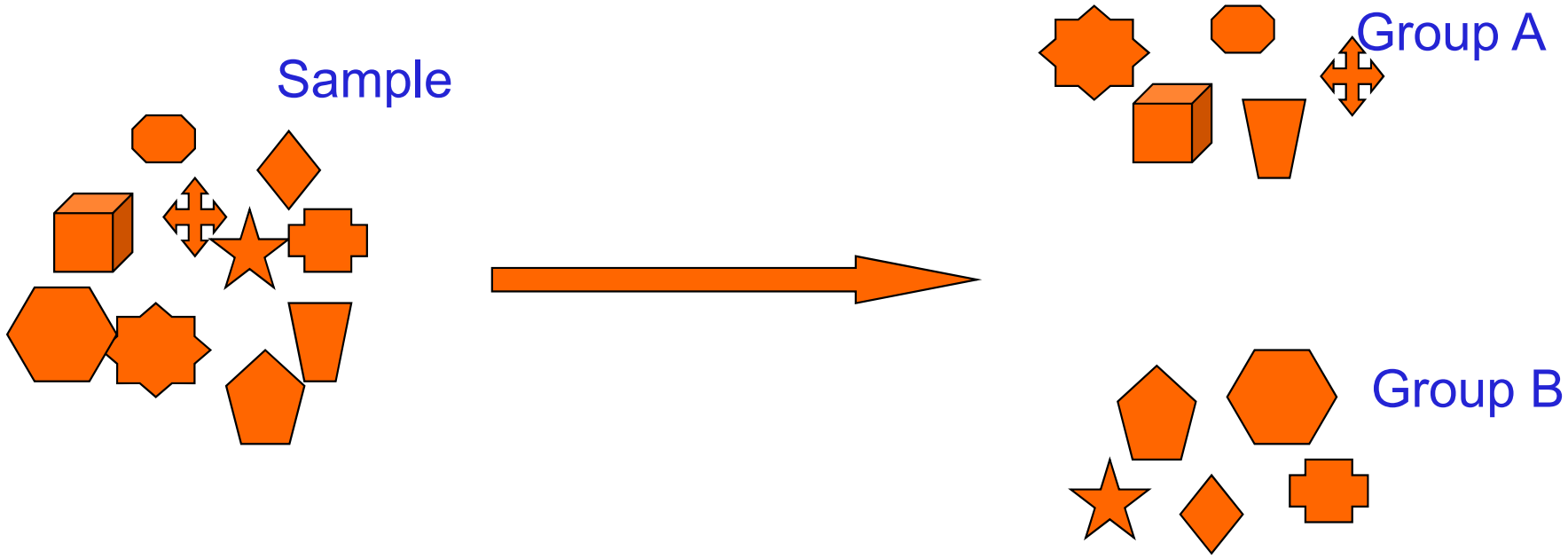
Random sampling vs. Randomization (Randomized Trial)

- Random sampling: to pick subjects from a population that are representative of that population
- Purpose: to estimate the truth in population from the facts in sample



Randomization

- Randomize to two or more groups:
  to minimize the difference of characteristics between groups (control selection bias)



Planning the Measurements:
Precision & Accuracy

Introduction

- Measurements => describe phenomena => analyzed statistically.
- Internal validity: how well the measurements represent these variables.
- External validity: how well the variables represent the phenomena of interest.

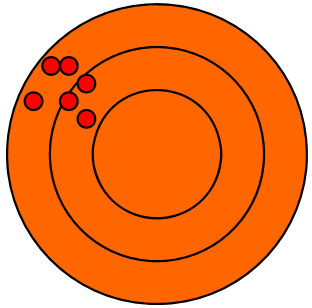
Measurement Scales

- Continuous variables:
  - Continuous variables: weigh, height, length,...
  - Discrete variables: a finite number of intervals, ex. number of cigarettes a day, age,
- Categorical variables:
  - Binominal variables = Dichotomous variables. Ex. Sex, death, ...
  - Nominal variables: unordered categories. Ex. blood type.
  - Ordinal variables: ordered categories, unquantifiable intervals. Ex degree of pain, severity of disease.

Choosing a Measurement Scale

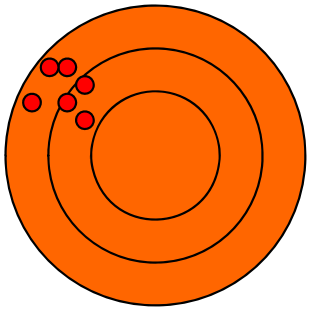
- A continuous variable can be collapse to a categorical variable (not vice verse).
- Ex. BP (mmHg) (discrete variable)
  => degree 0, I, II, III (ordinal variable)
  => Hypertension or Normal (binominal)

Precision



- Reproducibility, Reliability, Consistency
- More precise => the greater the statistical power (at the same sample size)
- Effected by random error
- Three main sources of error:
  - Observer variability
  - Subject variability
  - Instrument variability

Precision



- Assessing Precision: the consistency of repeated measurements.
  - Within-observer reproducibility
  - Between-observer reproducibility
  - Within-instrument reproducibility
  - Between-instrument reproducibility

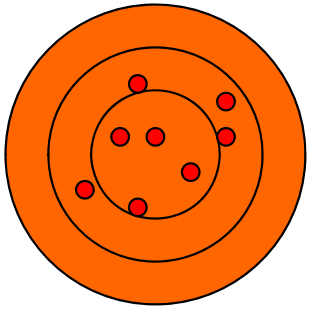
Strategies for Enhancing Precision

1. Standardizing the measurement methods
   - Operational definitions
   - Operation manual
2. Training & certifying the observers
3. Refining the instruments
4. Automating the instruments
5. Repetition

The decisions based on:

- Feasibility & cost of the strategy
- Importance of the variable
- Magnitude of the potential problem with precision

Accuracy



- The degree it actually represents what it is intended to represent.
- Important influence on the internal & external validity of the study.
- Accuracy =><= Systematic error (bias)
- Three main sources of error:
  - Observer bias
  - Subject bias
  - Instrument bias

Assessing Accuracy

- Comparing to "Gold Standard":
  - Continuous scale: compare means.
  - Categorical scale: sensitivity & specificity.
- No a Gold Standard available:

  => assess accuracy (validity of measurement)

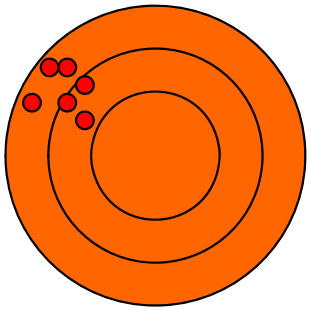
Strategies for Enhancing Accuracy

1. Standardizing the measurement methods
2. Training & certifying the observers
3. Refining the instruments
4. Automating the instruments
5. Making unobtrusive measures
6. Blinding (subjects, observers)
7. Calibrating the instruments

The decisions based on:

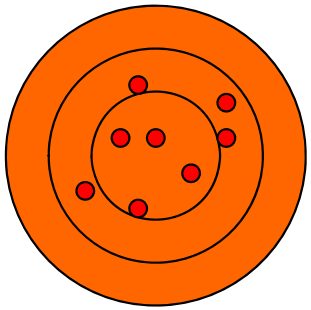
- Feasibility & cost of the strategy
- Importance of the variable
- Magnitude of the potential impact with inaccuracy

Precision - Accuracy



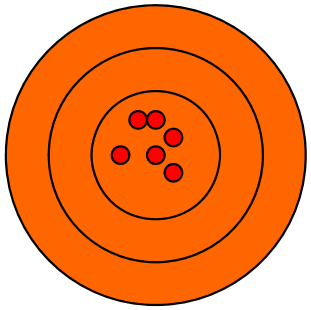
Good precision

Poor accuracy



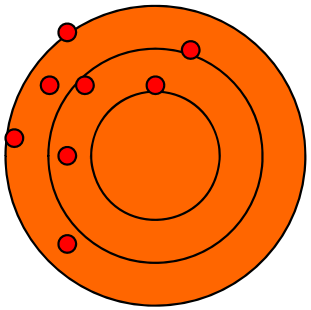
Poor precision

Good accuracy



Good precision

Good accuracy



Poor precision

Poor accuracy

Exercise 1

- Objective: to investigate the prevalence of metabolic syndrome in population >= 18 yo.
- Design: .....
- Target population: ....
- Sampling: how to recruit objects ....
- Measurements:
  - Hypertension: ....
  - Obesity: ....
  - Lipid profile: ....
  - Blood glucose ....

Exercise 2

- Objective: Is there relationship between metabolic syndrome and cardiovascular disease?

- Design: ....
- Objects: ....
- Measurements: (list all possible variables)
  - ......
  - ......
  - ......