PUBH 526

Topic 1, Part 2: Introduction and Review

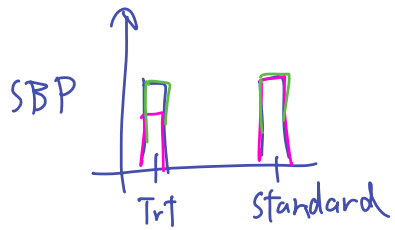
Prof. Nilupa Gunaratna

Department of Public Health

Let's Design a Study…

You want to test a new drug for its efficacy in reducing systolic blood pressure

- How would you design such a study? ★ RCT; Trt v.s. control; ★ Measure SBP after Trt.
- What would be your hypothesis? Trt lowering SBP compared to control
- What would data consistent with your hypothesis look like? ① Trt lowering SBP (favorable Effect) ② Same Effect
- What would data inconsistent with your hypothesis look like? Trt increasing SBP (not favorable)
- What kind of statistical test would you use? ① two-sample t-Test (comparing mean SBP between groups) ② ANOVA if >2 groups if adjusting for covariates



Design of Experiments

- Different disciplines have different applications, approaches, vocabulary, methods, history, literature
- Origins in agriculture
- Industry and engineering
- Health sciences
- Social sciences, including economics

Fundamental Principles

Randomization

Ensures fair comparisons by:
① distributing known/unknown factors across groups.
② ↑ validity of causal inference

- What is randomized? -- unit of randomization Ex) subjects, population, individuals, groups
  - Individual people
  - Classrooms at a school
  - Villages in a rural area
  - Counties in Indiana
- Measurement and analysis may occur at the unit of randomization or at higher or lower levels
  - Units of randomization (who/what), units of measurement/analysis (where outcomes are measured), context Ex) clustering...
- Why randomize?
  - Balance out effects of "lurking" variables – confounders, effect modifiers, mediators…
  - Many of these may be unknown
  - Does "balance out" mean "eliminate"? NO. Goal: Make Trt v.s. Control Comparable on average

Treatments

- In public health:
  - Treatment effects take time
  - The control treatment may not be obvious
- Counterfactual: what would have happened if there were no intervention?
  - Changes over time? – secular? seasonal? Ex) Measuring depression? Seasonal in winter
- Often difficult to separate the intended treatment from the delivery method
  - New drug vs. placebo
  - How about a health informational session?

Estimating Treatment Effects

- Accuracy is about hitting the target
- Precision is about spread

- What exactly is the target here? And what is an estimate?

Truth
Population Parameters → Targets

Estimate → Red Dots

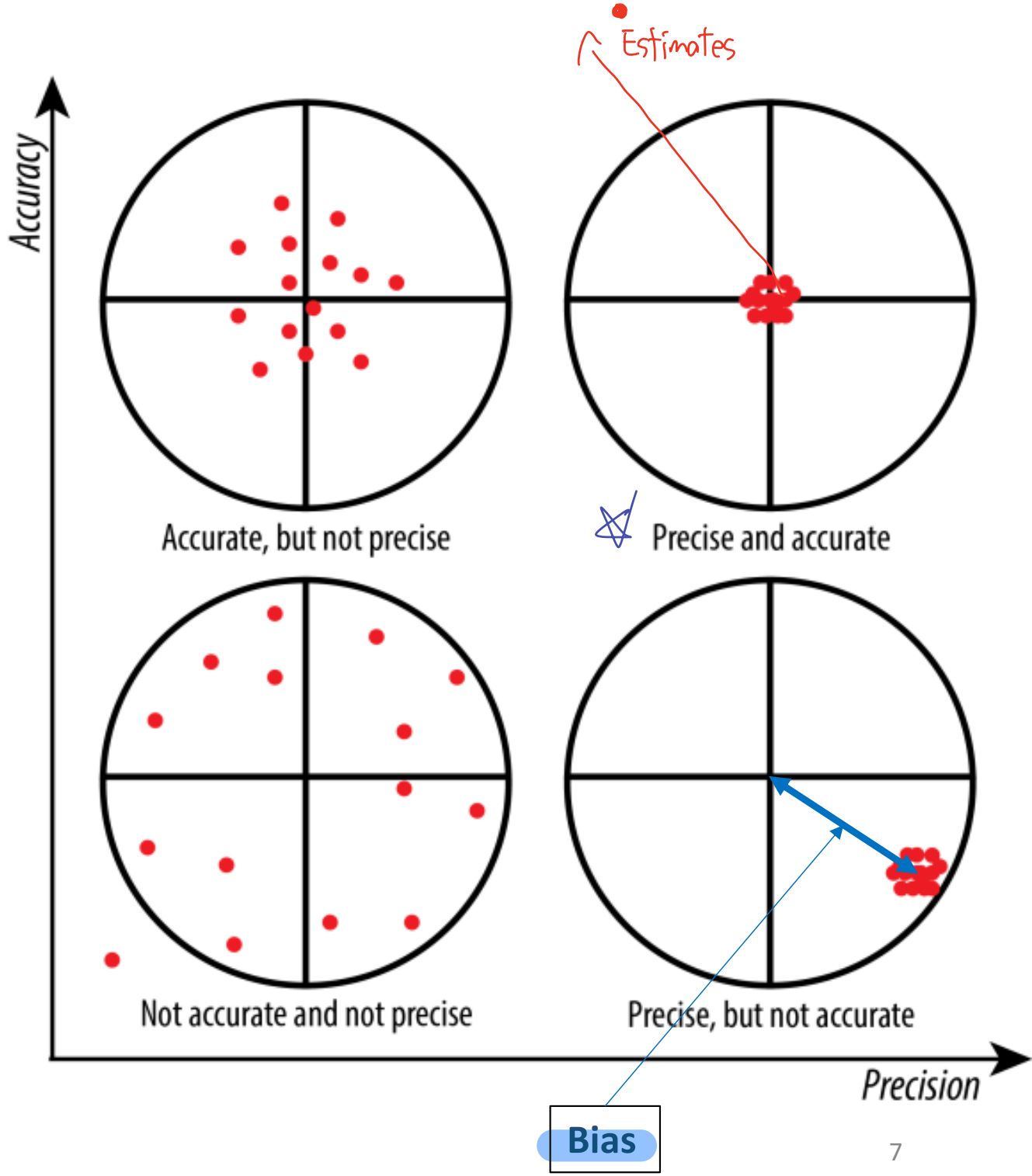
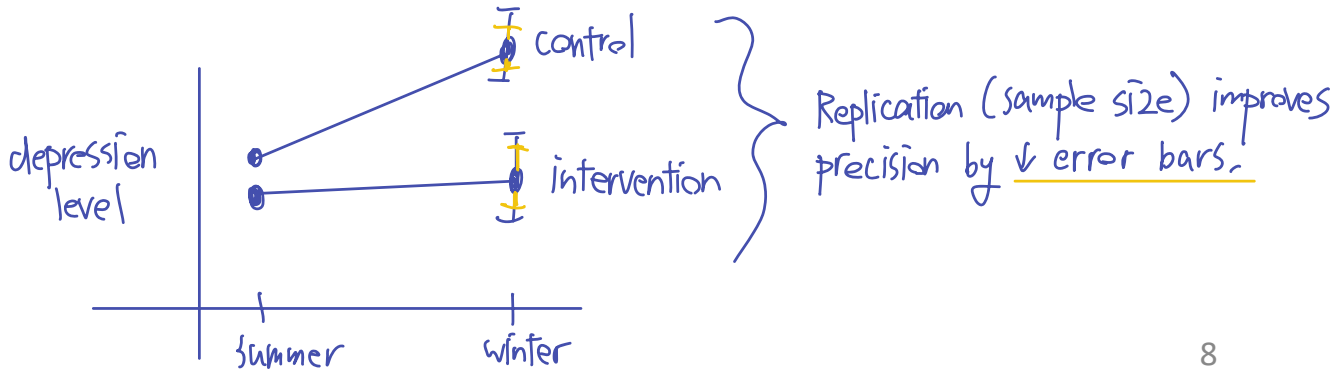


Image source: https://wp.stolaf.edu/it/gis-precision-accuracy/

o Replication ↑ Power

Replication

- Sample size ↑ probability of detecting a true effect if one exists (↓ Type 2 error: false neg.)
- Source of statistical power
- A greater sample size improves the precision ↑ of your ↓ CI estimates, including your estimate of the error or background noise ↑ stability, ↓ sensitivity to random fluctuations
- Consider sample size at all levels of measurement, particularly at the level of randomization



◦ Blocking ↑ precision

Blocking

- A way of dealing with "nuisance" factors
- Stratification
- Different blocks have different levels of the nuisance factor, but within a block, participants are more similar
- Examples?
  - **Agriculture:** Fields differ in soil fertility → block by field, randomize treatments within each field.
  - **Clinical trial:** Patients differ by sex or baseline health → block by sex/health category, randomize treatment within each block.
  - **Education:** Schools differ by size → block by school, randomize classrooms within each.

◦ Blinding ↓ bias

Blinding

- Why?
- Who?
  - Participants, researchers (many roles here), both (double blinding)
    - Don't know if they receive Trt or Control
    - Don't know which subjects are in which group.
    - Either patient nor doctor knows
- How? (blinded, unblinded, codes only)
  - Everyone knows the Trt. (when blinding is not possible)
  - Triple blinding (+ analysts, reviewers)
- When?
  - ↓ Concealed
  - Ex) Placebo looks same as Trt pills

Blinding

- How do you randomize while maintaining double blinding?
  - Roles: PI, senior personnel, pharmacists, physicians, nurses, data collectors, specimen collectors, data managers/cleaners/analysts, participants, … (blind at different levels at different times)
    - Trt must look and feel the same (identical placebo pills/procedures)
  - How to ensure consistency of treatments for the study duration?
    - Blinding must be broken if safety concern
  - How do you maintain participant safety?
  - Anticipate adverse outcomes, collect data equally in all study arms
- When is blinding not possible?
  - Of participants? Of researchers?

when is blinding not possible?

↳ Giving foods to participants (they know what the food looks like) — Researchers/staff

↳ People preparing/sharing foods

↳ Community Interventions (Messages conveyed)

↳ Lifestyle Interventions Ex) diet, exercise…

Participants

Internal and External Validity

- **Internal validity:** The degree to which the results are attributable to the **experimental treatment** and not some other explanation within study sample
- **External validity:** The extent to which the results of a study can be **generalized** outside the study

Feature	Internal Validity	External Validity
Concern	Is the causal relationship true *inside* the study?	Can results be applied *outside* the study?
Threatened by	Bias, confounding, measurement error	Lack of representativeness, artificial setting
Goal	Correct inference about cause & effect	Generalizability to real-world populations
Improved by	Randomization, blinding, replication	Diverse samples, pragmatic study design, replication across contexts

Internal Validity

IV comes first !
(EV)
- If your study isn't valid inside, you can't generalize it!

- Internal validity allows us to make statements about causality
  - However, mechanisms may not be clear
- Randomization aims to balance "explanations" across treatment arms
  - At start of study and before
  - During the study (assuming "explanations" affect both groups similarly)
  - Implementation should be balanced as much as possible – and monitored
- (Double) blinding -- to minimize biases of researchers and participants
- Attrition and missing data
  - Differential across treatments?
  - If attrition is the similar across treatments, is everything ok?

External Validity

- How representative is this sample of the larger population?
  - Did participants elect to participate?
  - What was the burden of participation?
  - Context
- Is the larger population clearly defined?
- How will the larger / other populations react to the specific treatment / intervention?
- Will the treatment remain the same when scaled up? How about the quality of implementation?

Designs to Be Covered in This Course

- Completely randomized design
- Randomized complete block design
- Crossover design
- Cluster randomized design
- Factorial design
- Other topics…