# Week 1 Review

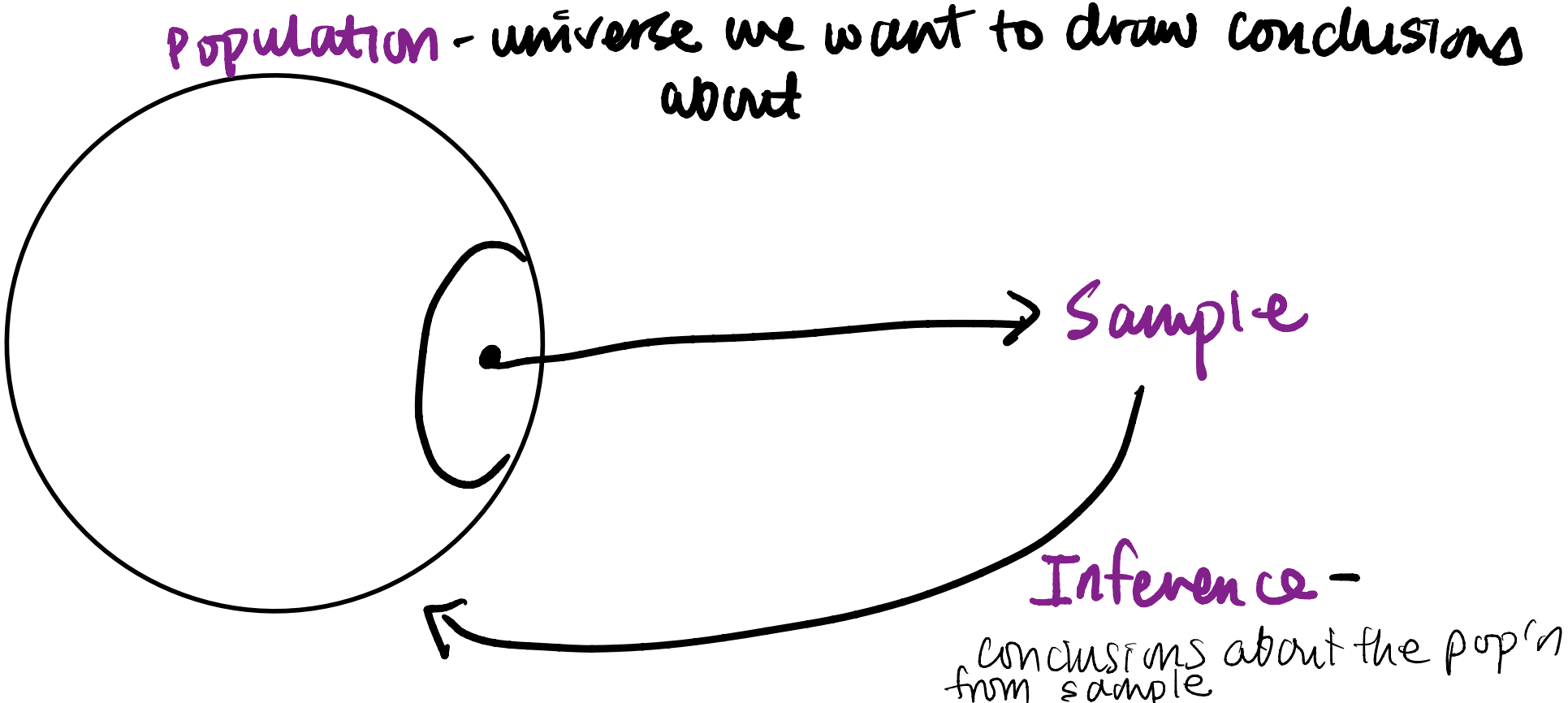
Cindy J. Pang

BIOSTAT 100A Summer Session C 2024

August 9, 2024

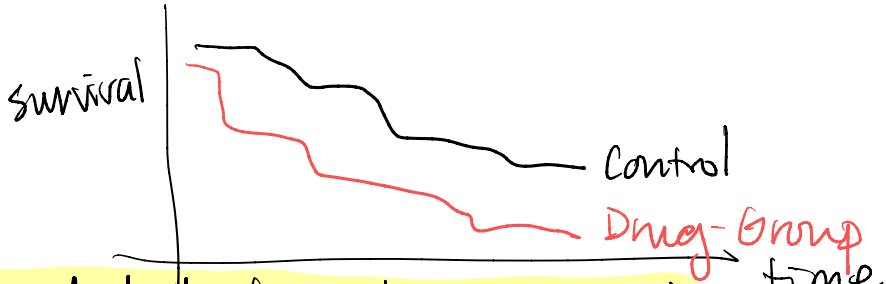
# Lecture 1: Introduction to Biostatistics

- Population vs. Sample
- Inference



# Lecture 1: Introduction to Biostatistics

- Problems Unique to Biostatisticians:
  - 1. Survival Analysis ~ time to event


censored data (right censoring)
↳ we never observe what happens after trial period.

  - 2. Clinical Trials (see exercise on next slide)
    - What is our **Population?** Everyone ≥16 y.o.
    - Who is in our **Sample?** Everyone who completed the study
  - 3. Survey Sampling (see exercise on next, next slide):
    - What is our **Population?** Everyone who lives in the United States
    - Who is in our **Sample?** People who complete the survey

# Exercise 1: Clinical Trial → did the vaccine work? → Hypothesis Testing Problem.


**BACKGROUND**

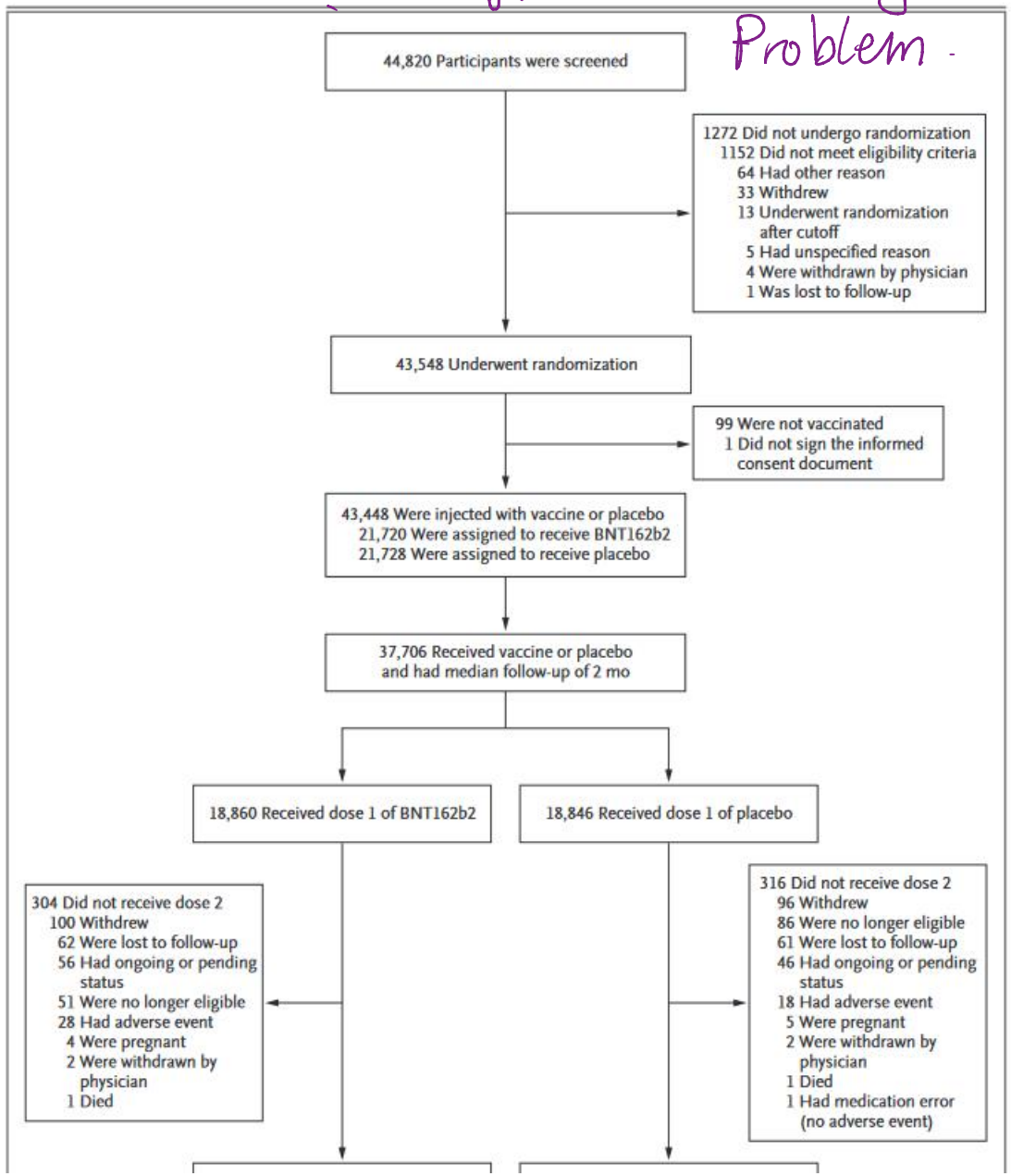
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and the resulting coronavirus disease 2019 (Covid-19) have afflicted tens of millions of people in a worldwide pandemic. Safe and effective vaccines are needed urgently.

**METHODS**

In an ongoing multinational, placebo-controlled, observer-blinded, pivotal efficacy trial, we randomly assigned persons 16 years of age or older in a 1:1 ratio to receive two doses, 21 days apart, of either placebo or the BNT162b2 vaccine candidate (30 μg per dose). BNT162b2 is a lipid nanoparticle–formulated, nucleoside-modified RNA vaccine that encodes a prefusion stabilized, membrane-anchored SARS-CoV-2 full-length spike protein. The primary end points were efficacy of the vaccine against laboratory-confirmed Covid-19 and safety.

**RESULTS**

A total of 43,548 participants underwent randomization, of whom 43,448 received injections: 21,720 with BNT162b2 and 21,728 with placebo. There were 8 cases of Covid-19 with onset at least 7 days after the second dose among participants assigned to receive BNT162b2 and 162 cases among those assigned to placebo; BNT162b2 was 95% effective in preventing Covid-19 (95% credible interval, 90.3 to 97.6). Similar vaccine efficacy (generally 90 to 100%) was observed across subgroups defined by age, sex, race, ethnicity, baseline body-mass index, and the presence of coexisting conditions. Among 10 cases of severe Covid-19 with onset after the first dose, 9 occurred in placebo recipients and 1 in a BNT162b2 recipient. The safety profile of BNT162b2 was characterized by short-term, mild-to-moderate pain at the injection site, fatigue, and headache. The incidence of serious adverse events was low and was similar in the vaccine and placebo groups.


From NEJM: **Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine**

# Exercise 2: Survey Sampling

The U.S. Census

Census → you sample the entire population

**Table 1. Demographic Analysis Estimates of Net Coverage Error in the 2020 Census for the Population Ages 0 to 4 by State** → Estimation Problem

| State FIPS Code | State Name | DA Population Estimate | 2020 Census Count | Net Coverage Error Estimate |
|---|---|---|---|---|
| 01 | Alabama | 297,751 | 286,529 | -3.77 |
| 02 | Alaska | 50,255 | 48,104 | -4.28 |
| 04 | Arizona | 419,488 | 392,370 | -6.46 |
| 05 | Arkansas | 189,309 | 179,575 | -5.14 |
| 06 | California | 2,319,173 | 2,137,439 | -7.84 |
| 08 | Colorado | 325,309 | 314,580 | -3.30 |
| 09 | Connecticut | 181,819 | 176,831 | -2.74 |
| 10 | Delaware | 54,992 | 51,230 | -6.84 |
| 11 | District of Columbia | 44,083 | 37,095 | -15.85 |
| 12 | Florida | 1,143,120 | 1,030,284 | -9.87 |
| 13 | Georgia | 651,900 | 614,218 | -5.78 |
| 15 | Hawaii | 85,659 | 77,352 | -9.70 |
| 16 | Idaho | 114,638 | 114,128 | -0.44 |
| 17 | Illinois | 738,282 | 705,616 | -4.42 |
| 18 | Indiana | 420,162 | 408,828 | -2.70 |
| 19 | Iowa | 195,743 | 190,064 | -2.90 |
| 20 | Kansas | 185,068 | 179,446 | -3.04 |
| 21 | Kentucky | 274,385 | 264,254 | -3.69 |

https://www.census.gov/data/tables/2020/demo/popest/2020-state-county-da-tables.html

# Lecture 2: Types of Data

- Data Taxonomy "quantifiable"

(1) [redacted] Quantitative - a meaningful number can be assigned.

ex: blood pressue, age, ♡ rate

- Continuous - any value is possible within a reasonable range; no gaps.

ex: time, (0,1), rainbow, $\mathbb{R}$ (real numbers)

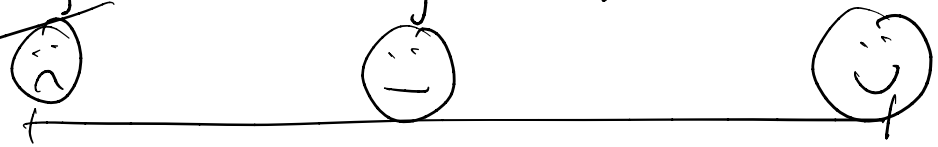
- Discrete - observations that can take a finite # of values

ex: Integers ($\mathbb{Z}$), $N = 0, 1, \cdots, n$, # classes UCLA Students take

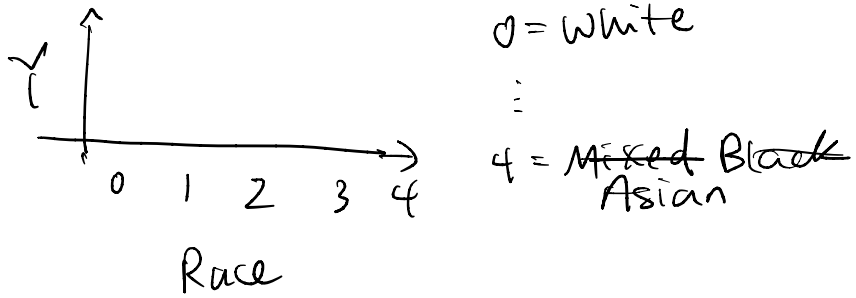
(2) [redacted] Qualitative - can't assign a meaningful #, categories

ex: Race, Hair Color, Mood (Gray Area)

~ Pain Scale Example



# Lecture 2: Types of Data




- Stevens' Scale of Data Classification

| | Categorical | Order | Add/Subtract | Mult./Divide | Examples |
|---|---|---|---|---|---|
| Nominal | X | | | | voting party, race |
| Ordinal | | X | | | pain scale, letter grades |
| Interval | | X | X | | time on a clock, Farenheit & Celsius scales (Temp.) |
| Ratio | | X | X | X | Kelvin Scale (0 = absence of energy) |

does not have a true Ø

has a true Ø ⇒ the absence of something.

# Lecture 3: Introduction to Sampling

- Rationale: want a sample representative of the population.
- Types of Samples:
  - Probability (Random) Sampling - every sample in the population has a chance of being selected
    i.e. P(choose $S_i > 0$) where $S_i$ is sample i
  - Non-Probability Sampling (*know examples of non-probability samples*) - not every sample has the chance of being selected
    (1) Convenience Sampling
    (2) Internet Sampling → people w/o internet access are systematically excluded
    (3) Clinical Research → b/c exclusion criteria e.g. pregnant woman, allergic to certain medications, mental health disorders, etc.)

| | When/Why do this type of sampling | How to conduct this type of sampling | How to select participants | Examples |
|---|---|---|---|---|
| Simple Random Sampling (SRS) - every sample has an equal chance of selection | When:<br>- Randomnization → reduces Bias<br>- Convenience<br>Assumption - assume respondents in sample are homogeneous → problematic | (1) Population List / Population Frame<br>(2) Assign a unique ID to each person in the frame<br>(3) Draw randomly. | | Lottery |
| Stratified Random Sampling | | | | |
| Cluster Random Sampling | | | | |
| Systematic Random Sampling | | | | |