

# A Statistician Reads the Newspaper

Howard Seltman

1. About me
2. Some key statistical principles
3. Examples
4. Review / lessons learned
5. Questions

# My Brief Bio

- 20 years in medicine (clinical pathology)
  - Medical College of PA
  - University of Pittsburgh
  - EVMS (Jones Institute for Reproductive Med.)
- 25 years in statistics
  - Carnegie Mellon University
- 2 years retirement

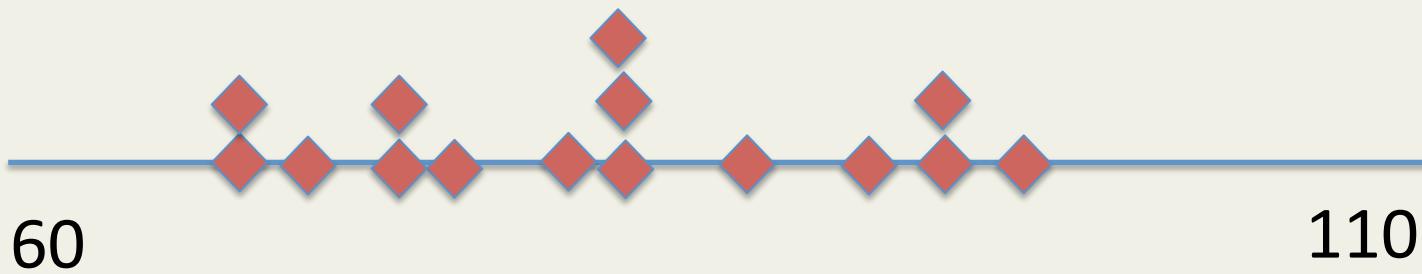
# Statistical Principles

- Rabbi Hillel was able to teach a potential convert to Judaism all of the Torah while standing on one foot: “What is hateful to you do not do unto your neighbor.”
- Today I can only choose a few important principles

# Principle 1: Randomness

- Making claims in the real world is hard due to randomness (variation)
- Important sources of randomness include:
  - subject-to subject variation
  - measurement variation
  - environmental variation
  - treatment application variation
- Example: masking for COVID 19 protection

# Randomness

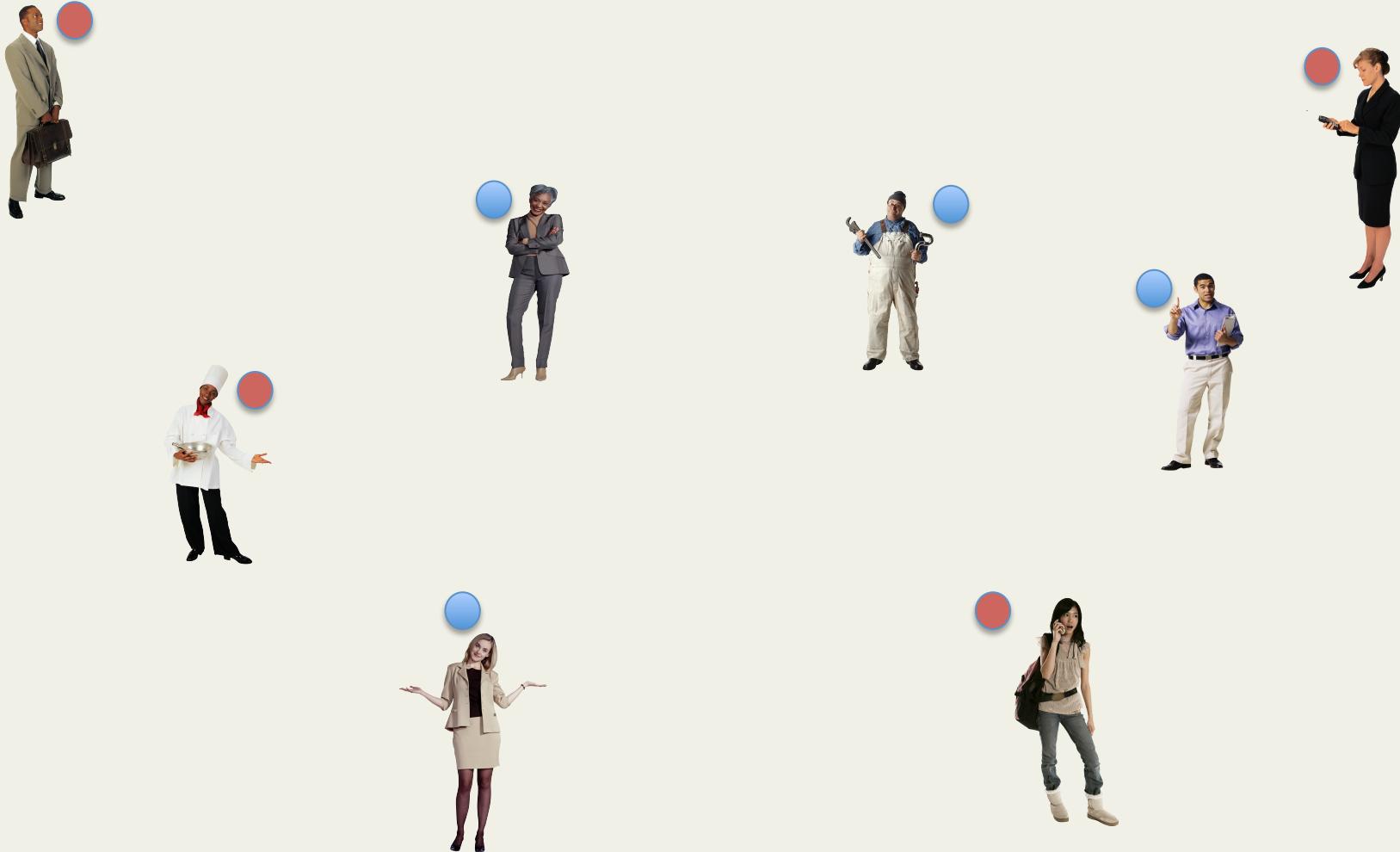


Diastolic Blood Pressure  
On Medication

# Principle 2: Compare Like Things

- Control and treatment group
- Only the treatment differs across groups
- Most effective method (by far): randomly assign treatment
- Example: Vitamin-C to prevent colds
- Additional consideration: blinding

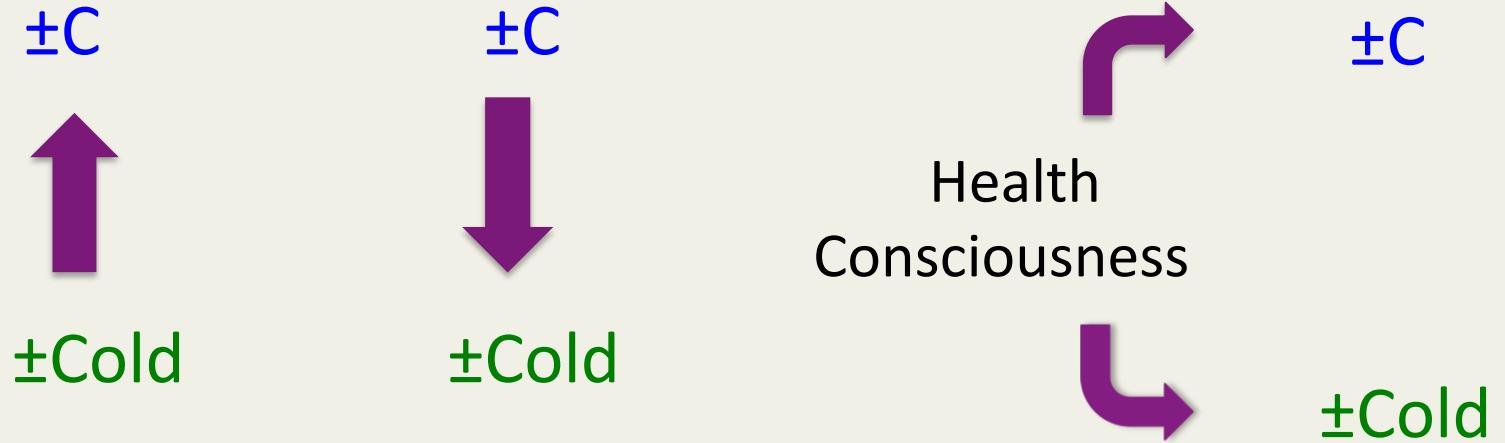
# Compare Like Things



# Principle 3: Causation is Tricky

- If groups differ in the average (mean) of some measurement, then either
  1. the difference is due to randomness
  2. the group “labels” are causing the change in what is measured
  3. what is measured is causing the values of the “labels”
  4. an external common cause is affecting both
- Examples
- Randomized experiments eliminate 3 and 4

# Causation Is Tricky

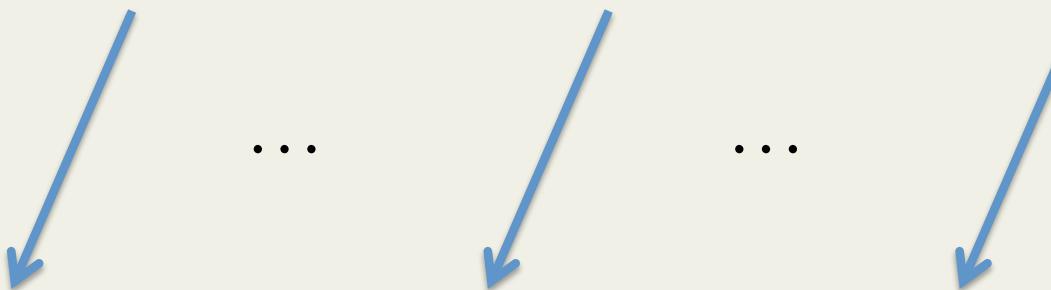


# Principle 4: The Classical Method

- (as opposed to Bayesian statistics)
- Imagine an ideal world: true population means of outcome for each group plus some amount of variation within groups
- For a given sample size, compute the distribution of the *possible* differences in sample means **if** the population difference is 0.
- Conclude the true means are different if the *observed* mean difference is larger than 95% of the *possible* differences when there is no “effect”

# Classical Method

Population



Samples



No drug

Drug

# Consequences of Classical Method

- If a study is repeated, the observed group mean difference in outcome will be different, and sometimes the conclusion will be different.
- If no true difference exists, the chance of falsely claiming a difference is 5%.
- If a true difference does exist, the chance of falsely concluding “no difference” varies with the power, which depends on the size of the true difference, the inherent variability, and the sample size
- 95% confidence intervals are a good (often better) alternative to yes/no conclusions.

# Consequences of Classical Method

- Terminology
  - Null hypothesis (no true difference)
  - Rejecting the null hypothesis (sample difference is larger than typically would be seen if the null hypothesis is true) is called statistically significant
  - Retaining the null hypothesis (sample difference is within the range commonly expected if the null hypothesis is true) means no good evidence for a claim of an “effect”
- Statistical significance is **not** the same as practical or scientific significance

# Example 1: COVID Vaccine

- **Morbidity and Mortality Weekly Report (CDC)**
- The Advisory Committee on Immunization Practices' Interim Recommendation for Use of Pfizer-BioNTech COVID-19 Vaccine in Children Aged 5–11 Years — United States, November 2021
- November 12, 2021 / 70(45);1579–1583
- [https://www.cdc.gov/mmwr/volumes/70/wr/mm7045e1.htm?s\\_cid=mm7045e1\\_w](https://www.cdc.gov/mmwr/volumes/70/wr/mm7045e1.htm?s_cid=mm7045e1_w)

# Example 1: Vaccine, cont.

- “The Pfizer-BioNTech COVID-19 vaccine has high efficacy (>90%) against COVID-19 in children aged 5–11 years, and benefits outweigh risks for vaccination. Vaccination is important to protect children against COVID-19 and reduce community transmission of SARS-CoV-2.”

# Example 1: Vaccine, cont.

- “The evidence ... was primarily composed of data from one randomized, double-blind, placebo-controlled phase II/III clinical trial that initially enrolled 2,268 participants aged 5–11 years, randomized 2:1 to receive vaccine or saline placebo.
- Interim findings from this clinical trial were based on data from participants with a median follow-up of 3.3 months.

# Example 1: Vaccine, cont.

- Vaccine efficacy was 90.9% (95% CI = 68.3%–98.3%) in preventing symptomatic, laboratory-confirmed COVID-19 in children aged 5–11 years ... based on infection in three vaccine recipients and 16 placebo recipients, none of whom were hospitalized.
- The measure of immune response ... was at least as high as the response observed in persons aged 16–25 years, with a ... mean ratio for ... antibody titer of 1.04 (95% CI = 0.93–1.18).

# Example 1: Vaccine, cont.

- “Serious adverse events were uncommon and occurred with similar frequency among vaccine (0.07%) and placebo (0.10%) recipients, with no statistically significant difference .. between the two groups.” [1 in each group]
- “An expanded safety cohort of 2,379 children (including 1,591 vaccine recipients) was added to monitor for serious adverse events, which had a median follow-up of 2.4 weeks.... No serious adverse events related to the vaccination were identified in either group.”
- **Rule of 3:** Rate is likely to be less than  $3/1591=0.2\%$

# Example 1: Vaccine, cont.

- What is good and bad in this study?

# Example 2: Bedtime

- Inc. Magazine, Oct. 13 2021, by Bill Murphy Jr.
- “A Massive New Study of 88,000 People Says This Is When You Should Go to Bed, According to Science”
- “It turns out, there's a magic hour. Here's why it matters.”
- <https://www.inc.com/bill-murphy-jr/>

## Example 2: Bedtime, cont.

- “[T]he UK Biobank, … [tracked] the health and habits of more than 500,000 British people over 30 years, beginning in 2006.”
- “Participants answered health questions, had basic measurements taken … and provided samples of blood and urine … [and some] wore sleep-tracking devices on their wrists for seven days.”

## Example 2: Bedtime, cont.

- “A group of UK scientists was able to take some of this data -- relating to about 88,000 participants -- and use it to track the correlation between sleep habits and propensity to develop heart disease during the five-and-a-half years that followed.”
- “The study was published in the journal [European Heart Journal -- Digital Health.](#)”
- “[T]hose who went to bed consistently between 10 p.m. and 11 p.m. had by far the lowest incidence of heart trouble”

## Example 2: Bedtime, cont.

- “Now, it's important to recognize that our old friend, ‘correlation versus causation,’ is a big issue here. We can't go so far as to say that going to bed between 10 p.m. and 11 p.m. actually causes a lower risk of heart trouble; it's just that people who went to bed at that time were less likely to develop heart issues for whatever reason.”

## Example 2: Bedtime, cont.

- Study author: "While we cannot conclude causation from our study, the results suggest that early or late bedtimes may be more likely to disrupt the body clock, with adverse consequences for [heart] health"
- Journalist: "I'll add another supposition, ... people who consistently go to bed at the eminently reasonable hour of 10 p.m. might also be more likely to practice other behaviors that more directly affect health..."

## Example 2: Bedtime, cont.

- Journalist's conclusion: "So if it works for you, maybe follow the science and focus on getting to bed not too long after 10 p.m. each night for a while."
- What are the strong and weak points of this study and its reporting?

# Example 3

Ars Technica (9/30/2021, by Beth Mole)

“A year later, 45% of COVID patients in Wuhan still have symptoms. Fatigue, sweating, chest tightness, anxiety and myalgia were most common.”

Source: **JAMA Netw Open**. 2021;4(9):e2127403, by Zhang, et al. (Third Military Medical University and University of London)

## Example 3: Post-COVID, cont.

- “This retrospective, multicenter cohort study was conducted at 2 designated hospitals [...] in Wuhan, China. All adult patients with COVID-19 discharged between February 12 and April 10, 2020, were screened for eligibility. Of a consecutive sample of 3988 discharged patients, 1555 were excluded (796 declined to participate and 759 were unable to be contacted) and the remaining 2433 patients were enrolled. All patients were interviewed via telephone from March 1 to March 20, 2021.”

# Example 3: Post-COVID, cont.

- Subjects: half are men; 28% had severe COVID; median age is 60
- Partial results:
  - 45% has at least one symptom at 1 year follow up
  - 27% had fatigue
  - Higher risk of fatigue is associated with female sex ( $OR=1.27$ , 95% CI 1.06 to 1.52) and severe disease ( $OR=1.43$ , 95% CI 1.18 to 1.74).
  - 1.27 OR here is equal to 25% vs 29%
  - 2.43 OR here is equal to 22% vs. 28%

# Example 3: Post-COVID, cont.

- Similar results are given for other symptoms: sweating (17%), chest tightness (13%), anxiety (10%), and muscle ache (8%).
- Some less common symptoms had a larger odds ratio for severe vs. not-severe, e.g., impaired smell (OR 2.6, 2.5% vs. 0.9%) and sore throat (OR 3.1, 1.8% vs. 0.7%).
- What is good and bad in this study?

# Example 4: Diet

- [Today.com](#) (Today Show) Nov. 11, 2021, by Kristin Kirkpatrick
- “A diet low in these 4 foods could increase risk of dementia, study finds”
- “The observational study found that a diet low in fruits, vegetables, beans and tea was linked [sic] to three times greater risk of dementia.”

## Example 4: Diet, cont.

- “The study, published in the journal [Neurology](#), the medical journal of the American Academy of Neurology, showed that a diet low in fruits, vegetables, beans and tea was linked [sic] to three times greater risk of dementia.”
- \$39 paywall

## Example 4: Diet, cont.

- “The observational study was conducted in 1,059 Greek individuals with an average age of 73 and over a three-year period.”
- “Participants ... were assessed a score associated with inflammatory factors associated with their food choices. Researchers found that the higher the score, the higher the risk for dementia.”
- “The study was observational, meaning that the takeaway of the data can only show that there is an association (but not definitive proof) that anti-inflammatory diets may play a role in the prevention of dementia.”

## Example 4: Diet, cont.

- “Advanced age, genetics and ethnicity can all play a role in the risk of dementia. They are not modifiable, meaning, they cannot be changed. The good news, however is this: another major factor associated with dementia is one you can control. Your diet.”
- “Numerous studies have shown that lifestyle choices play a large role in brain health. How you approach your diet is the first step.”

## Example 4: Diet, cont.

- Multiple dietary choices are suggested
- “Protecting our brains as we age is doable. Get more nutritional bang for your buck by choosing foods that are nutrient dense — and give your brain the TLC it deserves.”
- What is good and bad in this study?

# Review and Lessons Learned

- Pay close attention to the study characteristics
  - ✧ Randomized vs. not is most important
  - ✧ Types of subjects
  - ✧ Outcomes
  - ✧ Size (bigger is better only for good studies)
- Look for reporter bias or overstatement

# Questions?

howard.seltman@gmail.com



# Example 5

- Insider (Andrea Michelson, Oct. 12, 2021)
- “Chemicals in shampoo and makeup are linked to early death, study finds”
- Source: “Phthalates and attributable mortality: A population-based longitudinal cohort study and cost analysis”, by Trasande, et al. (NYU **Environmental Pollution**, paywalled for \$35.95 at <https://doi.org/10.1016/j.envpol.2021.118021>)

# Example 5: Phthalates, cont.

- “Exposure to a common class of chemicals called phthalates was linked to premature deaths, especially deaths caused by heart disease, a new study ... found.”
- “Trasande and colleagues at NYU Grossman School of Medicine analyzed data from more than 5,000 adults between the ages of 55 and 64. They found that those with the highest level of phthalates in their urine were more likely to die earlier than expected, especially of heart-related causes.”

# Example 5: Phthalates, cont.

- “The study calculated that between 91,000 to 107,000 American adults a year suffer premature deaths linked to these chemicals. The authors put a price tag on those early deaths: an estimated \$40 billion to \$47 billion each year in lost economic productivity.”
- “The American Chemistry Council … sent CNN a statement calling the study ‘demonstrably inaccurate’ because it lumped all phthalates into a single group rather than considering the differences in toxicity.”

## Example 5: Phthalates, cont.

Design: 5303 adults aged 20 years or older who participated in the US NHANES Survey 2001–2010 and provided urine samples for phthalate metabolite measurements.

Participants were linked to mortality data [via death certificates] from survey date through December 31, 2015, recording death from any cause, cardiovascular disease, and cancer.

# Example 5: Phthalates, cont.

- Results:
  - High-molecular weight (HMW) phthalate metabolites: Hazard Ratio for overall mortality: 1.14 (95% CI 1.06-1.23)
  - di-2-ethylhexylphthalate (DEHP): HR for overall mortality 1.10 (95% CI 1.03-1.19)
  - Third third vs. first third: OR 1.48 (HMW) and 1.42 (DEHP)
  - Cardiovascular mortality was significantly increased for DEHP

# Example 5: Phthalates, cont.

- Conclusions:

“In a nationally representative sample, phthalate exposures were associated with all-cause and cardiovascular mortality, with societal costs approximating \$39 billion/year or more. While further studies are needed to corroborate observations and identify mechanisms, regulatory action is urgently needed.”

- What’s good and bad in this study?