

# Conceptualizing experimental controls using the potential outcomes framework

---

**IQSS Applied Statistics Seminar**

Kristen Hunter<sup>1</sup>, Kristen Koenig<sup>2</sup>, Marie-Abèle Bind<sup>3</sup>

April 21, 2021

<sup>1</sup> Harvard Department of Statistics

<sup>2</sup> Harvard Department of Organismic and Evolutionary Biology

<sup>3</sup> Massachusetts General Hospital Biostatistics Center

## A few extra pounds lowers mortality?

Meta-analysis of the relationship between weight and mortality found that being overweight, but not obese, was associated with a lower mortality than being normal or underweight (Flegal et al. 2013).

- Contradicted medical consensus that being overweight leads to poor health outcomes! (The Global BMI Mortality Collaboration 2016).
- Investigation: normal and underweight population contained excessive smokers, elderly, or chronically ill.



Photo by Vidmir Raic on Pixabay

## What went wrong?

This is an example of a poorly controlled study.

- Even with best intentions!
- Failed to isolate the effect of weight on mortality.
- Goal of a well-controlled study: the intervention of interest is the sole causal mechanism influencing the outcome.
- Could this study have been saved with an experimental control?

# Outline

1. Introduction to problem
2. Understanding experimental controls
3. Motivating example
4. Defining controls & applying controls
5. Application: air pollution and autism

# **Understanding experimental controls**

## Checks to improve study design

What is an experimental control?

- Framework for collecting additional information, leading to complementary calculations or analyses, to improve study design.
- An experimental control is an intentional set of experimental conditions, such as treatment levels and outcomes, that encode a scientific assumption.

Properties of controls:

- Based on prior knowledge.
- Designed to detect systematic (unwanted) variation.

## Applying a control to the obesity study

- *Control assumption:* weight has **no effect** on the rate of deaths due to motor vehicle accidents (Alban et al. 2006).
- *Additional analysis:* Compare deaths due to motor vehicle accidents in normal/underweight group to overweight group.
- *What we learn:* Rate of deaths by motor vehicle accidents is **lower** in normal/underweight group than overweight group.



## Applying a control to the obesity study

- Investigation: Difference in groups is caused by lower rates of motor vehicle deaths among elderly.
- Discovery: Elderly are **overrepresented** in the normal/underweight group.
- *Implication for main analysis:* We have **unwanted variation** in mortality due to something besides obesity.

# Objectives of this talk

1. Introduce and **understand** experimental controls found in natural sciences.
  - Outline clear, mathematically precise **definitions** of different experimental controls using the potential outcomes framework.
2. Advocate for experimental controls as useful **tools** in a statistical toolbox.
  - Provide guidance for the application of controls to **study design**.

## Applications of experimental controls

Experimental controls are a general framework for assisting with better study design, and can help with a wide variety of different challenges.

Preview: a few examples!

- Diagnosing unwanted variation.
- Determining optimal timing.
- Identifying responders.

## Relationship to other work

Related work and inspiration:

- Experimental design in **natural sciences** (Glass 2014).
- Negative and positive controls in **epidemiology** (Lipsitch et al. 2010; Shi et al. 2020).
- Placebo tests in **political science** (Hartman and Hidalgo 2018).
- Controls in **genetics** (Gagnon-Bartsch and Speed 2012).

Contribution of this work:

- Provide a **unifying framework** for defining controls.
- Translate concepts into the **potential outcomes** framework for clarity and accessibility.

## Motivating example

## Motivating (hypothetical) example

What is the short-term effect of drinking caffeinated coffee compared to drinking decaffeinated coffee on change in blood pressure? (Glass 2014)



Photo by Gordon Johnson on Pixabay

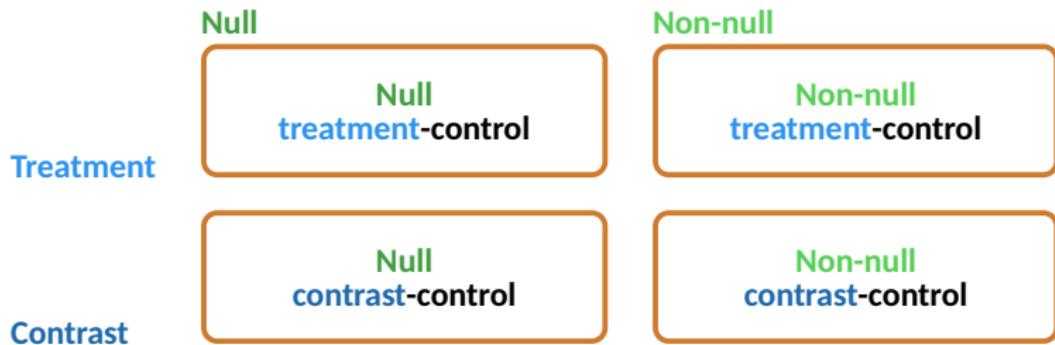
# Potential outcomes in caffeine experiment

- Treatment:  $W_i$ 
  - Active treatment level ( $w_{at}$ ): caffeinated coffee.
  - Control treatment level ( $w_{ct}$ ): decaf coffee.
- Outcome:  $Y_i^p$  (primary outcome)
  - Change in blood pressure before and after intervention.
  - Intervention occurs at time  $t$ :  $Y_i^p = BP_{i,t+1} - BP_{i,t-1}$ .
- Estimand: Difference in changes
  - Primary estimand:  $\tau_i^p = Y_i^p(w_{at}) - Y_i^p(w_{ct})$

Note: We will define other treatment levels and outcomes later.

## **Defining controls**

# Road map: defining types of experimental controls



## Treatment controls

## Null treatment-control

A **null treatment-control** is a treatment level  $w_{nt}$  for which the primary outcome  $Y_i^p$  is zero for unit  $i$ .

$$Y_i^p(w_{nt}) = 0.$$

## Null treatment-control: Caffeine example

## Null treatment-control: Caffeine example



- **No intervention** is a null treatment-control on change in blood pressure.

Photo by Rabie Madaci on Unsplash

**Null treatment-control**



**Non-null treatment-control**

**Null contrast-control**

**Non-null contrast-control**

## Non-null treatment-control

A **non-null treatment-control** is a treatment level  $w_{nnt}$  for which the outcome of interest  $Y_i^p$  is nonzero for unit  $i$ .

$$Y_i^p(w_{nnt}) \neq 0$$

## Non-null treatment-control: Caffeine example

## Non-null treatment-control: Caffeine example



- Taking a **hypertensive drug** (i.e. a drug that raises blood pressure) is a non-null treatment-control on change in blood pressure.

Photo by Sharon McCutcheon on Unsplash

**Null treatment-control**



**Non-null treatment-control**



**Null contrast-control**

**Non-null contrast-control**

## Practical example: Unwanted variation

During caffeine experiment, people sit in different waiting rooms.

Room 1: Great British Bake Off



Room 2: election results



Photo from Element 5 on Pixabay

Can result in **unwanted variation**: people in election results waiting room may have blood pressure increase **unrelated to caffeine**.

## Diagnosing unwanted variation

**Diagnostic tool:** null treatment-control

- Additional group: main interventions (active treatment, control treatment), and null treatment-control group (no intervention).
- Simple check: calculate average change in blood pressure in the null treatment-control group.
- This average change in blood pressure should be (centered around) zero.
- A **departure** from zero average blood pressure change is a sign of unwanted variation.

## Two consequences of unwanted variation

### Reduced precision

- If waiting room assignment is unrelated to treatment assignment, results are still unbiased.
- But would still increase the variance, so would reduce precision.

### Confounding

- If waiting room assignment is related to treatment assignment, would result in bias!
- Example: separate machines to make caffeinated and decaffeinated coffee.

## Practical example: Determining optimal timing

Q: How long should I wait to measure blood pressure after giving someone coffee?

A: Run a pilot study!

- Use a non-null treatment control (hypertensive drug).
- Measure how long it takes for hypertensive drug to increase blood pressure.
- Assumes that non-null treatment control and treatments of interest operate on similar time scales.



Photo by Ewa Urban on Pixabay

## **Control treatments**

## Control treatments

Active treatment:

- $[w_{at}]$  Caffeinated coffee

Possible control treatments:

- $[w_{ct1}]$  Water
- $[w_{ct2}]$  Caffeinated water
- $[w_{ct3}]$  Decaffeinated coffee

- Standard statistical definition of control: **control group**.
- Defined on a **single treatment level**.
- Experimental controls are a fundamentally different concept!

## Compared to what?

- What is the effect of drinking caffeinated coffee on blood pressure, **compared to what?** (Feller et al. 2016)
- We use control treatments (which are **not** necessarily null/non-null treatment-controls) to make our causal estimand **well-defined.**
- Our choice of control treatment determines our estimand  $Y(w_{at}) - Y(w_{ct})$ .

## Possible causal estimands

Control treatment	Estimand: $Y(w_{at}) - Y(w_{ct})$
$w_{ntc}$ : No intervention	Effect of all components of coffee
$w_{ct1}$ : Water	Effect of non-fluid component of coffee
$w_{ct2}$ : Caffeinated water	Effect of non-caffeine component of coffee
$w_{ct3}$ : Decaffeinated coffee	Effect of caffeine component of coffee

**Null treatment-control**



**Non-null treatment-control**



**Null contrast-control**

**Non-null contrast-control**

## **Contrast controls**

## Null contrast-control

A **null contrast-control** is an effect  $\tau_i^{nc}$  which is zero for unit  $i$ . It is defined as a contrast of potential outcomes between an active treatment level  $w_{atc}$  and a control treatment level  $w_{ctc}$ .

$$\tau_i^{nc} = Y_i^{nc}(w_{atc}) - Y_i^{nc}(w_{ctc}) = 0$$

## Null contrast-control: Caffeine example

## Null contrast-control: Caffeine example



- The effect of caffeinated coffee compared to decaffeinated coffee on **short-term change in body weight** is a null contrast-control.

Photo by i yunmai on Unsplash

**Null treatment-control**



**Non-null treatment-control**



**Null contrast-control**



**Non-null contrast-control**

## Non-null contrast-control

A **non-null contrast-control** is an effect  $\tau_i^{nnc}$  which is nonzero for unit  $i$ .

$$\tau_i^{nnc} = Y_i^{nnc}(w_{atc}) - Y_i^{nnc}(w_{ctc}) \neq 0$$

## Non-null contrast-control: Caffeine example

## Non-null contrast-control: Caffeine example

- The effect of caffeinated coffee compared to decaffeinated coffee on **change in reaction time** is a non-null contrast control (Santos et al. 2014).



Photo by Peggy und Marco Lachmann-Anke on  
Pixabay

## Practical example: Identifying responders

Q: What if some subjects don't respond to caffeine?

A: Collect data in a pre-trial period on a non-null contrast-control.

- Assumption: Define **non-responder** as someone who has **no physiological response to caffeine**
- Non-null contrast-control with treatments of interest (caf/decaf coffee) and secondary outcome known to be affected by caffeine (reaction time)
- If no effects for secondary outcome(s), person is a non-responder, and will not show effects in blood pressure either.

**Null treatment-control**



**Non-null treatment-control**



**Null contrast-control**



**Non-null contrast-control**



## **Application: air pollution and autism**

## Application: air pollution & autism

### Question

- Do higher levels of **air pollution** during pregnancy lead to a higher incidence of **autism**?
- Several studies have found an association; others have not.
- Such studies are very prone to **confounding**.

### Potential outcomes

- $Y_i^p(W_i)$ : primary outcome, autism (binary)
- $W_i$ : primary intervention, air pollution (binary: 1 if in top 25% of air pollution level during pregnancy, 0 if in bottom 25%)

**Disclaimer:** For illustrative purposes only!

# Applying controls to air pollution & autism

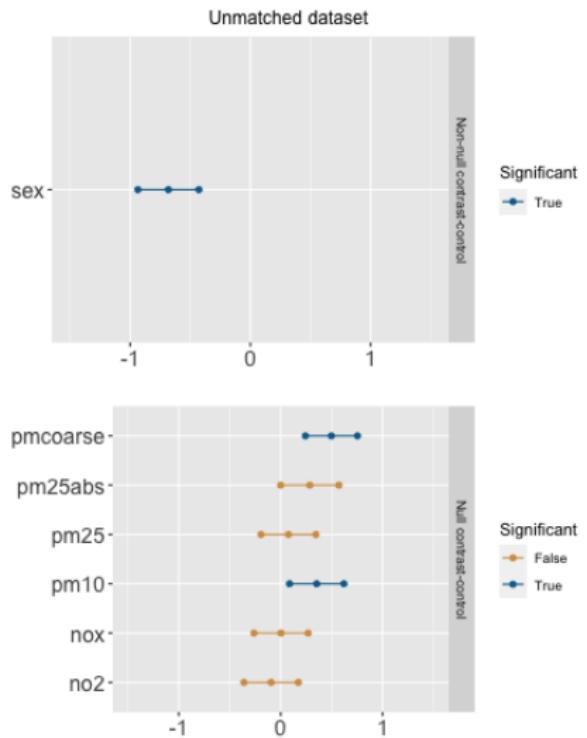
## Non-null contrast-control

- *Outcome:*  $Y_i^p$ , primary, autism
- *Intervention:* secondary, sex of child,  $w_{ats}$  = female,  $w_{cts}$  = male
- *Assumption:*  $Y_i^p(w_{ats}) - Y_i^p(w_{cts}) \neq 0$
- Nonzero causal effect of child's sex on autism

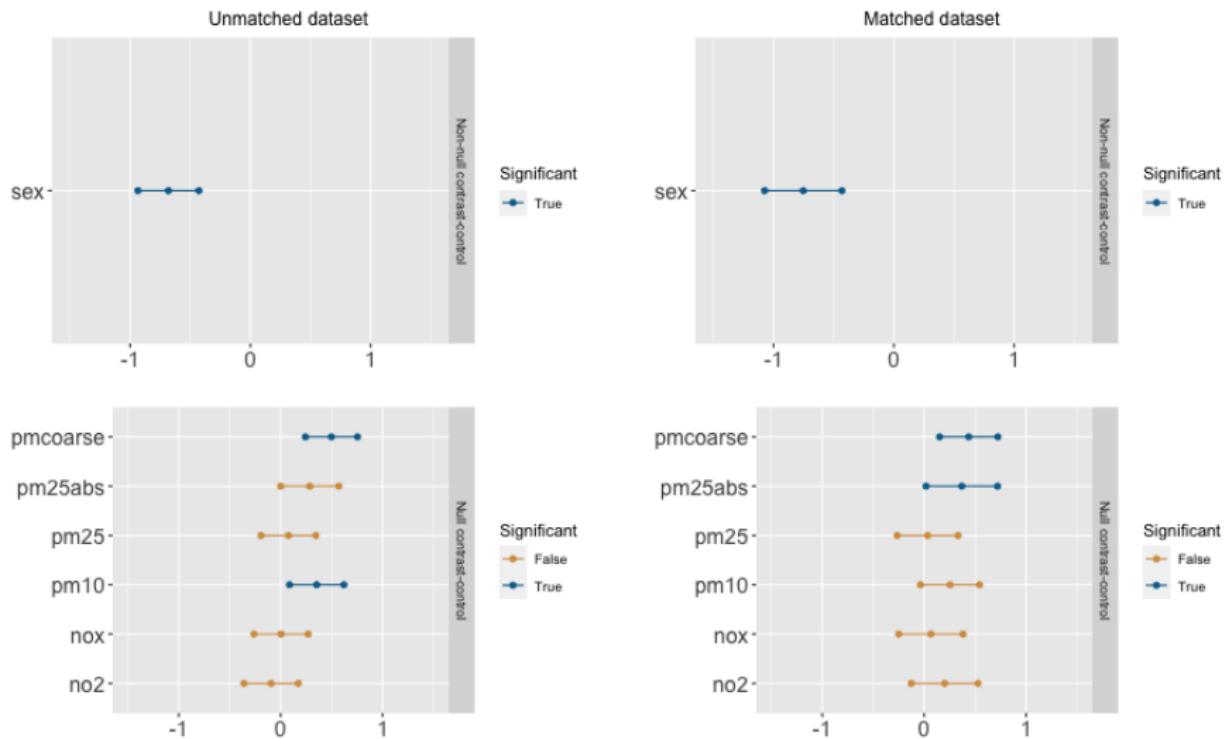
## Null contrast-control

- *Outcome:*  $Y_i^{nc}$ , secondary, maternal age at birth
- *Intervention:* primary, prenatal air pollution levels,  $w_{at}$ ,  $w_{ct}$
- *Assumption:*  $Y_i^{nc}(w_{at}) - Y_i^{nc}(w_{ct}) = 0$
- No causal effect of prenatal pollution on maternal age

# Control checks



# Control checks



# Conclusion

## Discussion: Applying controls to study design

- Tools to understand sources of variation.
- Controls can be useful at different stages of a study: pilot study, pre-trial phase, main phase.
- If a control analysis/test fails, we have gained insight into a particular failing of the experiment. If it does not fail, we merely have no evidence that any flaw occurred.
- Ideally, controls should be chosen to target potential areas of concern.

# Conclusion

Experimental controls are a framework for collecting additional information that can assist in improving study design.

- Potential outcomes framework to define controls on two axes:
  - null and non-null controls.
  - treatment and contrast controls.
- Practical applications of controls:
  - Diagnosing unintended factors
  - Determining optimal timing
  - Identifying responders

# Thank you!

Note: paper draft coming later today to arXiv!

## Acknowledgments:

- Bind Lab, CARES Lab, Gagnon-Bartsch Lab, Emma Thomas
- Stephane Shao, Alice Sommer, Terry Speed
- Autism dataset: Mònica Guxens

Research reported in this presentation was supported by the John Harvard Distinguished Science Fellows Program within the FAS Division of Science of Harvard University and by the Office of the Director, National Institutes of Health under Award Number DP5OD021412. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Kristen Hunter was also supported by the Department of Defense through the National Defense Science & Engineering Graduate Fellowship Program.

## References |

- Alban, R. F. et al. (2006). Obesity does not affect mortality after trauma. *The American Surgeon*, 72(10), 966–969.
- Feller, A., Grindal, T., Miratrix, L., & Page, L. C. (2016). Compared to what? Variation in the impacts of early childhood education by alternative care type. *The Annals of Applied Statistics*, 10(3), 1245–1285.
- Flegal, K. M. et al. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories. *Journal of the American Medical Association*, 309(1), 71–82.
- Gagnon-Bartsch, J. A., & Speed, T. P. (2012). Using control genes to correct for unwanted variation in microarray data. *Biostatistics*, 13(3), 539–552.
- Glass, D. J. (2014). *Experimental design for biologists* (2nd). Cold Spring Harbor Laboratory Press.
- Hartman, E., & Hidalgo, F. D. (2018). An equivalence approach to balance and placebo tests. *American Journal of Political Science*, 62(4), 1000–1013.
- Lipsitch, M., Tchetgen Tchetgen, E., & Cohen, T. (2010). Negative controls: A tool for detecting confounding and bias in observational studies. *Epidemiology*, 21, 383–388.

## References II

- Miao, W., & Tchetgen Tchetgen, E. (2018). A confounding bridge approach for double negative control inference on causal effect [https://arxiv.org/abs/1808.04945]. *arXiv:1808.04945*.
- Santos, V. G. et al. (2014). Caffeine reduces reaction time and improves performance in simulated-contest of Taekwondo. *Nutrients*, 6(2), 637–649.
- Shi, X., Miao, W., & Tchetgen Tchetgen, E. J. (2020). A selective review of negative control methods in epidemiology. *Current Epidemiology Reports*, 7, 190–202.
- The Global BMI Mortality Collaboration. (2016). Body-mass index and all-cause mortality: Individual- participant-data meta-analysis of 239 prospective studies in four continents. *The Lancet*, 388, 776–86.

## Defining controls

The term **control** is used to mean many different concepts.



- Statisticians often use control to mean a **control group**, such as giving one group a placebo treatment.
- Defined on a **single treatment level**.
- Control group: Essential to defining a causal estimand.

Photo from little plant on Unsplash

## Connection: Negative control outcomes and exposures

- A negative control outcome is a special case of a null contrast-control when the outcome is a secondary outcome and the treatment levels are the active and control treatment levels of interest (Lipsitch et al. 2010; Miao and Tchetgen Tchetgen 2018; Shi et al. 2020).
- A negative control exposure is a special case of a null contrast-control when the outcome is the primary outcome and the treatment levels are alternative active and control treatment levels.
- Analogous connections for positive control outcomes and exposures.

## **Outcome controls**

## Null outcome-control

A **null outcome-control** is an outcome  $Y_i^{no}$  that is zero for unit  $i$  given the active treatment level  $w_{at}$ .

$$Y_i^{no}(w_{at}) = 0$$

## Null outcome-control: Caffeine example



- Change in body flexibility is a null outcome-control for drinking caffeinated coffee.

Photo by Katee Lue on Unsplash

## Non-null outcome-control

A **non-null outcome-control** is an outcome  $Y_i^{nno}$  that is nonzero for unit  $i$  given the active treatment level  $w_{at}$ .

$$Y_i^{nno}(w_{at}) \neq 0$$

## Non-null outcome-control: Caffeine example

- Change in electrolyte concentration is a non-null outcome-control for drinking caffeinated coffee.



Photo by Gorkhs on Pixabay

## Practical example: Identifying compliers

Q: What if some subjects don't drink the coffee, and we don't know about it?

A: Collect additional information during the study on a non-null outcome control.

- Use a non-null outcome-control (electrolyte concentration)
- If electrolyte concentration does not change, the subject probably did not drink the coffee!
- Can possibly **identify** the important subgroup of **compliers**