When results defy common sense: A gentle entry to Bayesian methods

CKD 12

CKD 34

CKD 5

Hypertension = HT

PA = Physical Activity
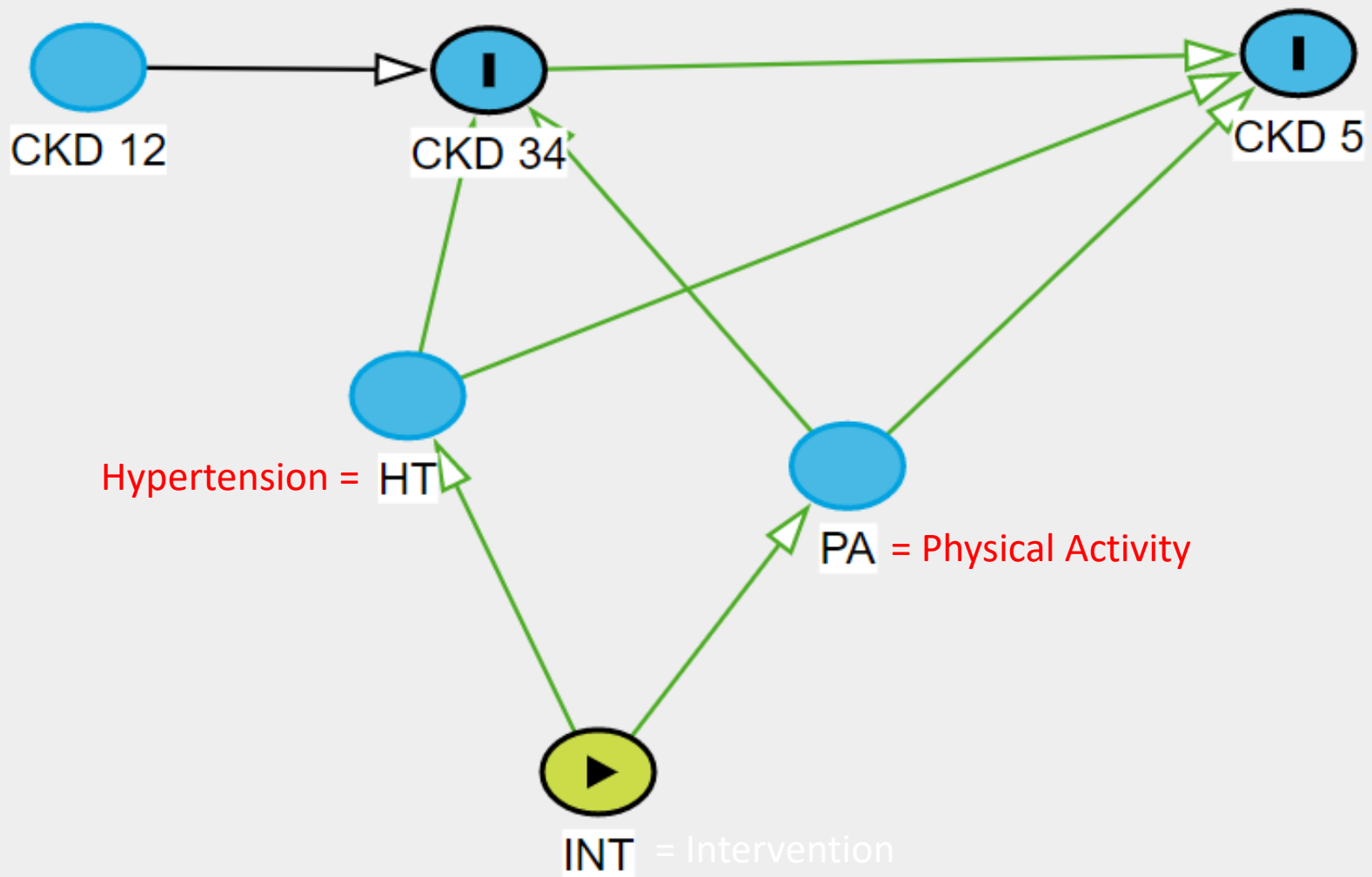
INT = Intervention

Estimation "intervention → mediator"

	Improvement in **hypertension**
Intervention	OR= 4.03 [1.14 to 14.31]

SRI

🔒 Restricted access | Research article | First published online February 14, 2011

Beautiful British Parents Have More Daughters

Satoshi Kanazawa View all authors and affiliations

≣ Contents | 🔒 Get access | 🔗 Cite article | � Share options | ⓘ Information, rights a

Abstract

The generalized Trivers-Willard hypothesis proposes that parents who possess any heritable trait

Follow up study surveyed 3000 Americans to test correlation between parents attractiveness sex of children.
56% of the children of parents in the highest attractiveness category were girls, compared to 48% of the children of parents in the other categories.
Difference of 8% (standard error 3%)

- Gertler (2013) "Labor returns to Childhood stimulation"
- Randomized stimulation to stunted Jamaican toddlers living in poverty

- Gertler (2013) "Labor returns to Childhood stimulation"
- Randomized stimulation to stunted Jamaican toddlers living in poverty
- One-hour weekly visits from community Jamaican health workers over a 2-year period. Earnings measured 20 years later (n = 127)

- Gertler (2013) "Labor returns to Childhood stimulation"
- Randomized stimulation to stunted Jamaican toddlers living in poverty
- One-hour weekly visits from community Jamaican health workers over a 2-year period. Earnings measured 20 years later (n = 127)
- Estimated increase in earnings 42%, Confidence interval [2%, 98%]

- Gertler (2013) "Labor returns to Childhood stimulation"
- Randomized stimulation to stunted Jamaican toddlers living in poverty
- One-hour weekly visits from community Jamaican health workers over a 2-year period. Earnings measured 20 years later (n = 127)
- Estimated increase in earnings 42%, Confidence interval [2%, 98%]

"We report substantial effects on the earnings of participants in a randomized intervention conducted in 1986–1987 that gave psychosocial stimulation to growth-stunted Jamaican toddlers.... the intervention had a large and statistically significant effect on earnings.... The estimated impacts are substantially larger than the impacts reported for the US–based interventions, suggesting that ECD interventions may be an especially effective strategy for improving long-term outcomes of disadvantaged children in developing countries."

Effect estimation - 42% CI [2%, 98%] (N=127, std error ≈ 0.12)

Suppose that the real effect on earnings is 10%
- What is the probability that we get significant results?
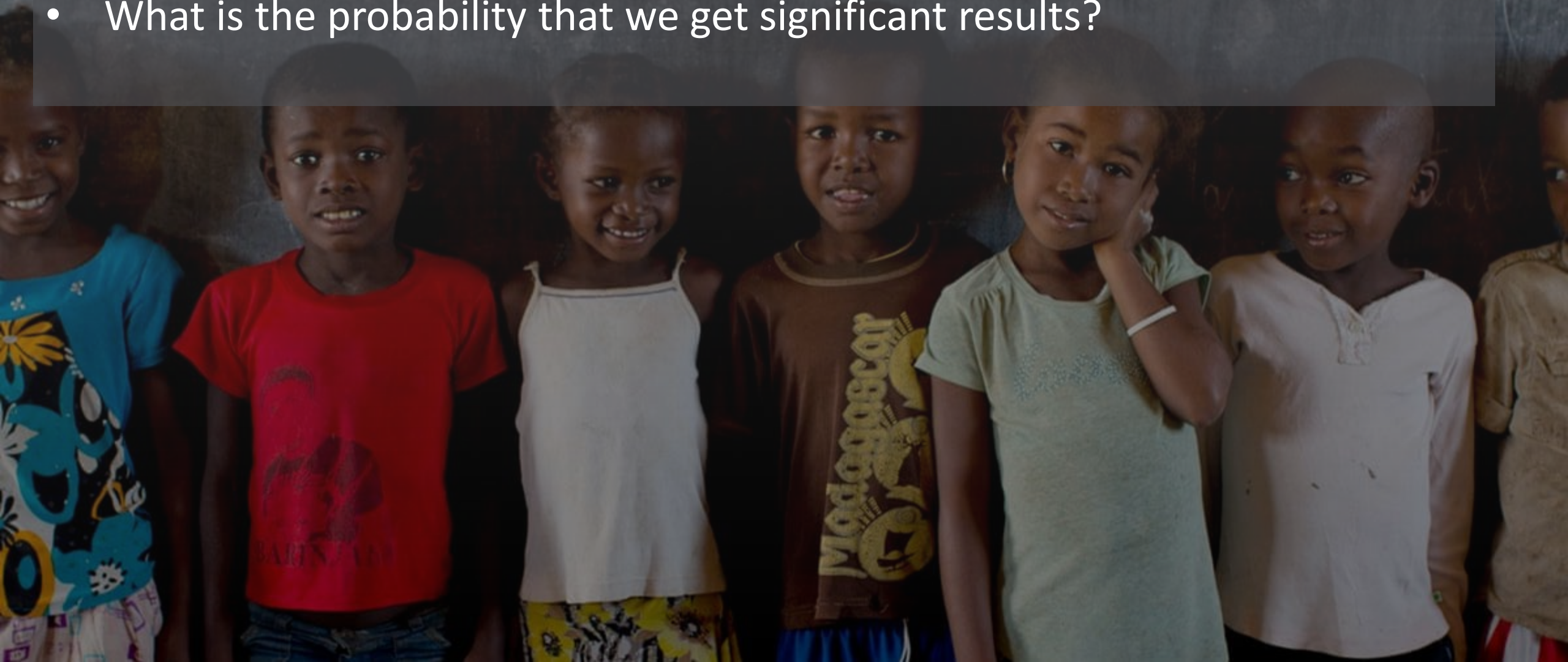
Effect estimation - 42% CI [2%, 98%] (N=127, std error ≈ 0.12)

Suppose that the real effect on earnings is 10%
- What is the probability that we get significant results?

```
significant?          n exp.mn

FALSE               892    1.07
TRUE                108    1.52
```
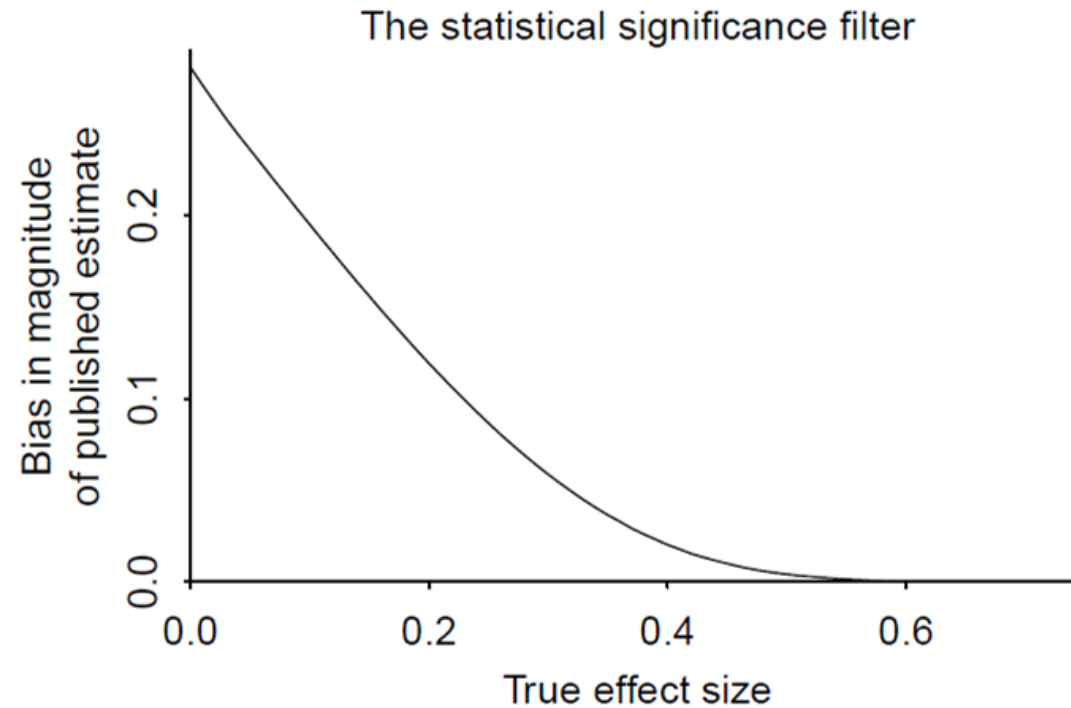
Effect estimation - 42% CI [2%, 98%] (N=127)

Figure 1: *Bias in expected magnitude of effect size estimate, conditional on statistical significance, as a function of actual effect size, for the early-childhood intervention study of Gertler et al. (2014). The raw estimate, before selection is assumed to be normally distributed with mean equal to the true effect and standard error 0.12.*

Frequentists: "Let the data speak for itself"

1 Hypothesis H_0

2 H_0: What should the data look like?

3 Test the data

4 Reject/Don't reject

Reverend
Thomas Bayes (1701 – 1761)

Bayesians: "Update your belief!"

1 Prior belief
2 Data
3 Updated belief

Bayesians: "Update your belief!"

1 Prior belief $\sim N(0, 0.10)$
2 Data 1.42 CI [1.02, 1.98]
3 Updated belief: **prior + data**
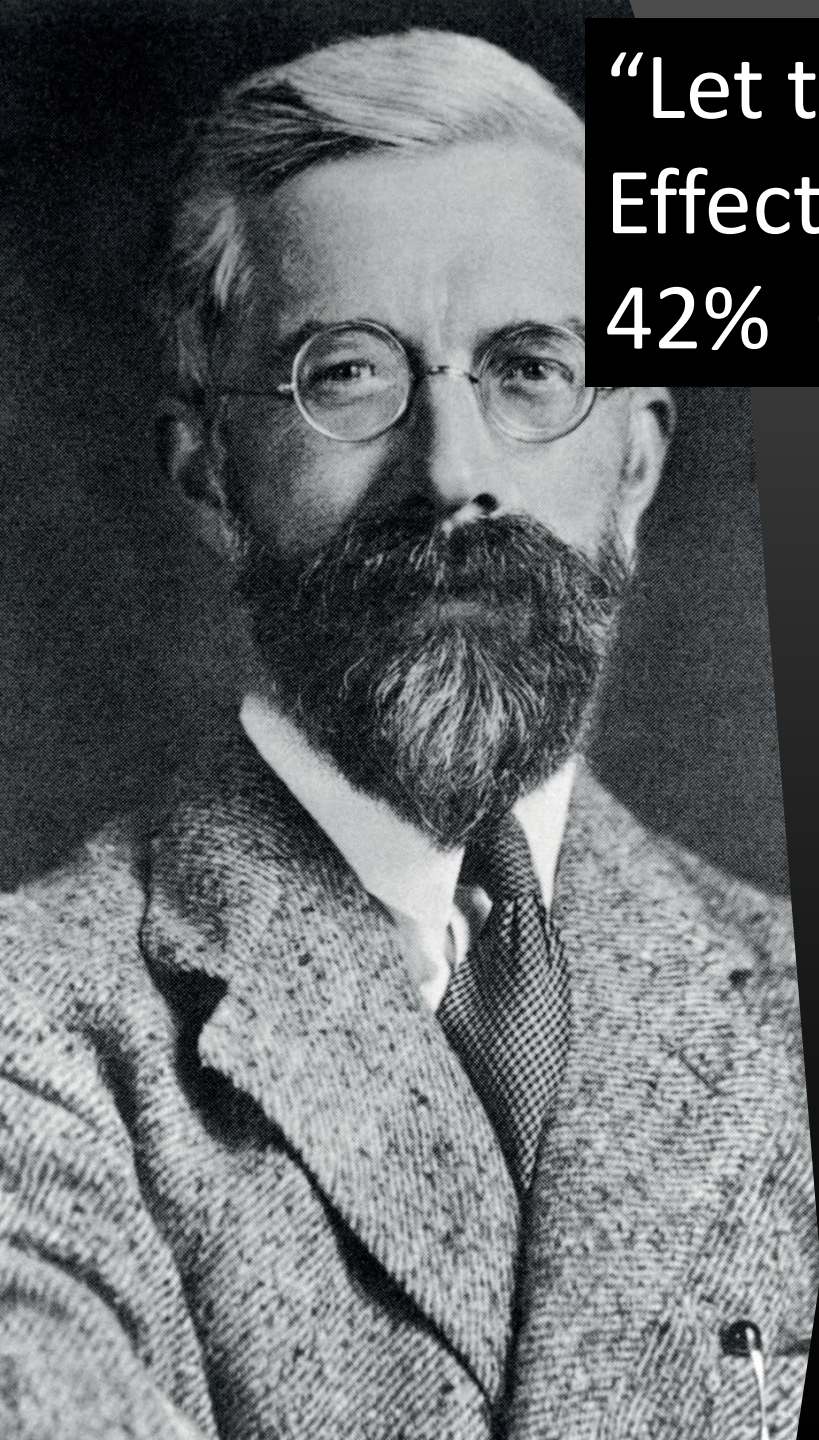
Bayesians: "Update your belief!"

1 Prior belief $\sim N(0, 0.10)$
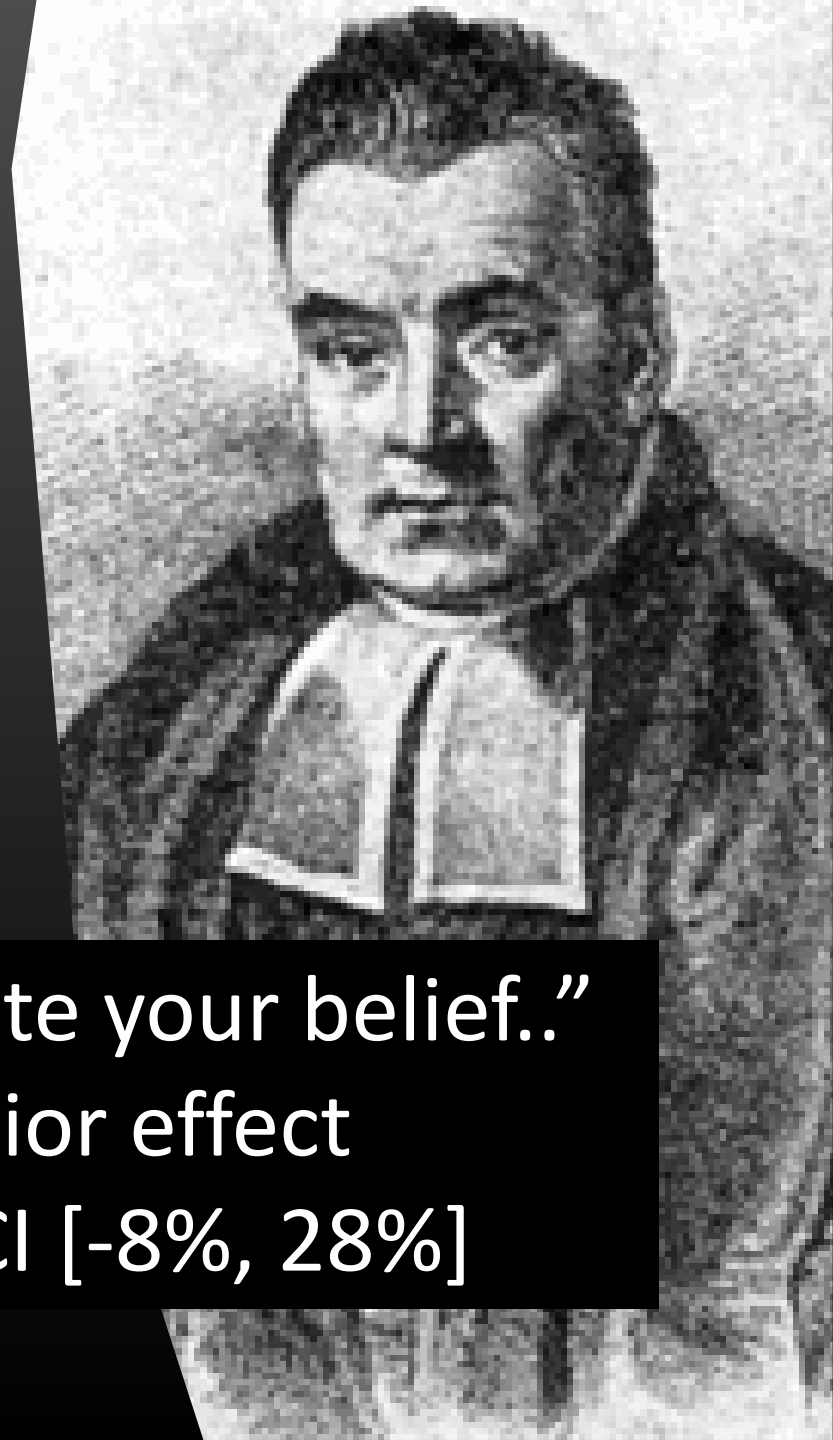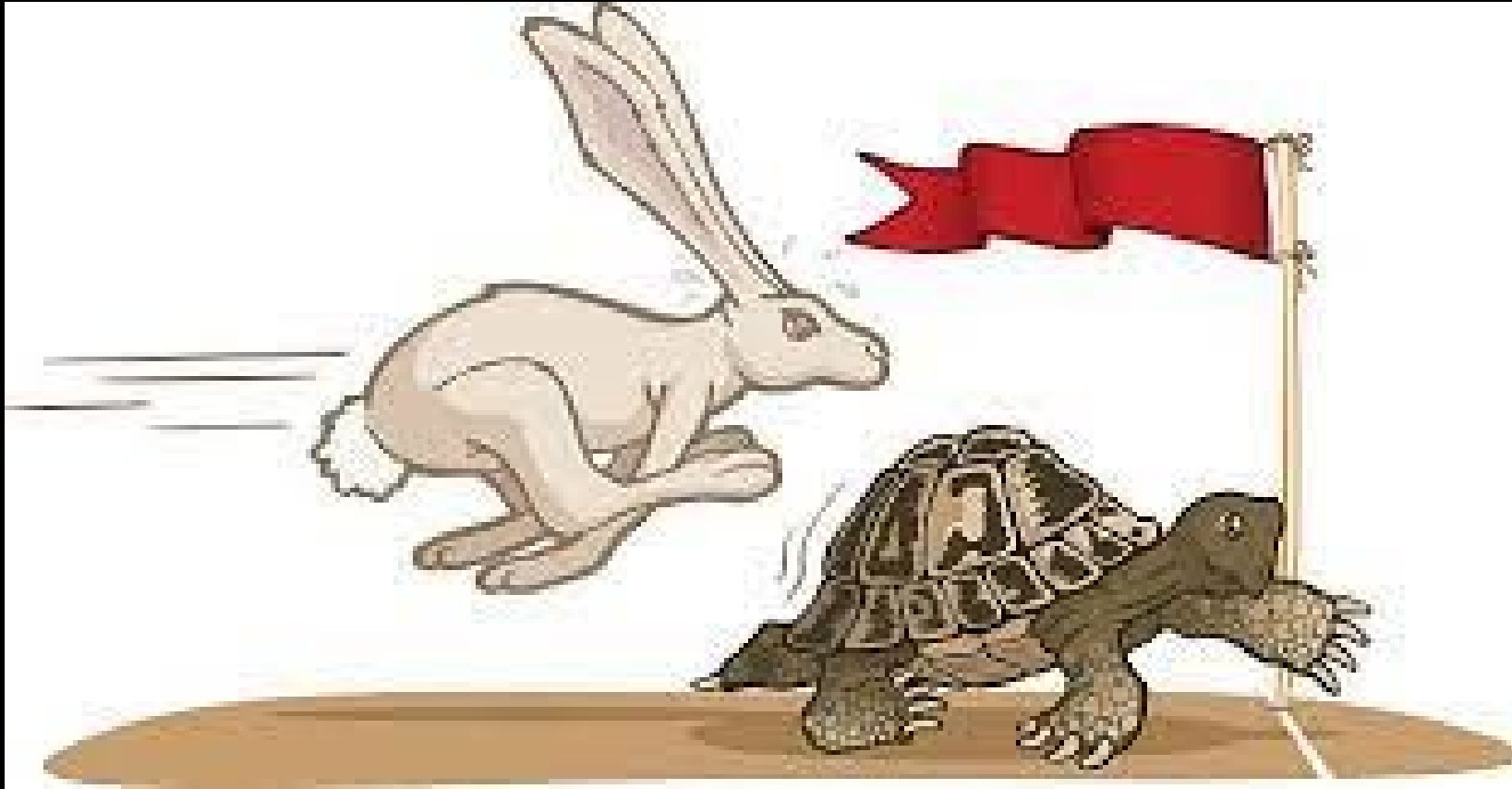2 Data 1.42 CI [1.02, 1.98]
3 Updated belief:

1.09 CI [0.92, 1.28]

"Let the data speak.."
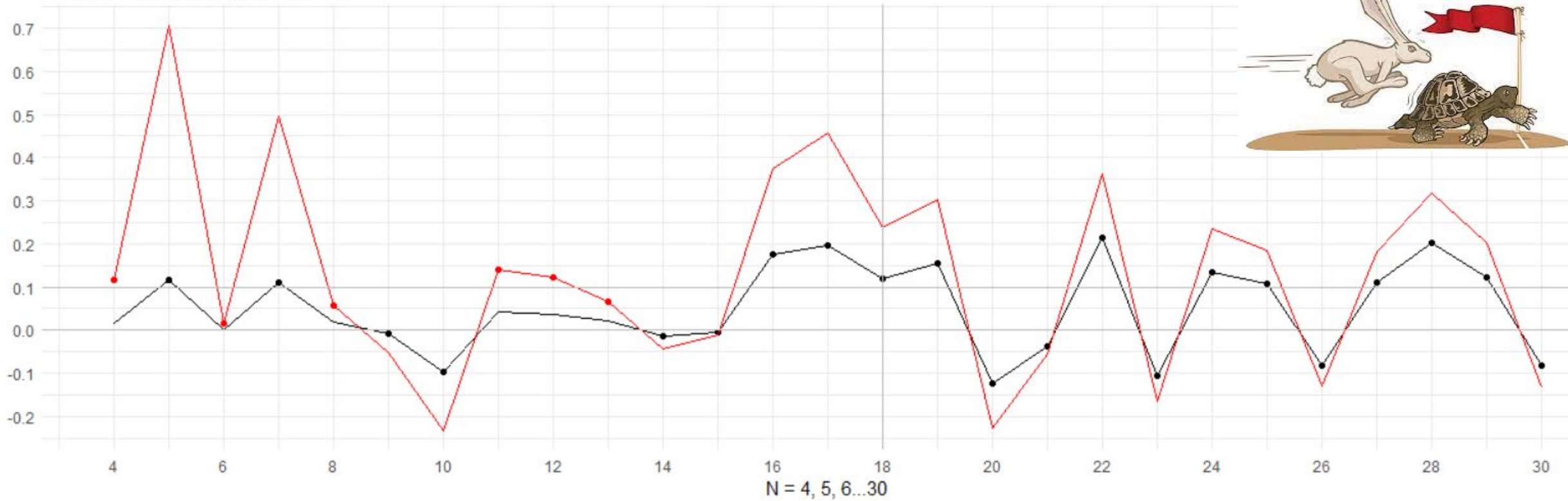Effect estimation:
42% CI [2%, 98%]

"Update your belief.."
Posterior effect
9%   CI [-8%, 28%]

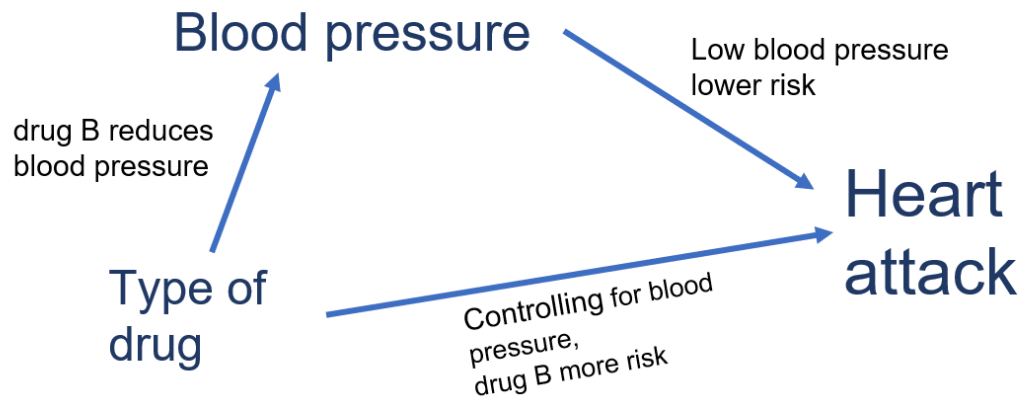$$y = 1 + 0.1 \cdot x + \text{error}$$

Estimating the slope
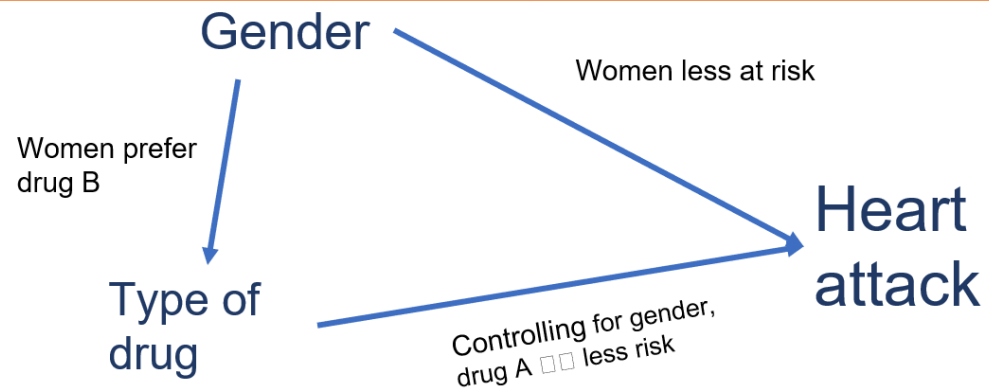Frequentist (red) vs Bayesian (black)

N = 4, 5, 6...30

$$y = 1 + 0.1 \cdot x + \text{error}$$
$$\text{error} \sim N(0, 0.5)$$

Top-left table

	Drug A			Drug B		
	Heart attack	No heart attack	% of patients with heart attacks	Heart attack	No heart attack	% of patients with heart attacks
Low blood pressure	1	19	5.0%	3	37	7.5%
High blood pressure	12	28	30.0%	8	12	40.0%
Total	13	47	21.7%	11	49	18.3%

Blood pressure

drug B reduces blood pressure

Low blood pressure lower risk

Type of drug

Controlling for blood pressure, drug B more risk

Heart attack

Bottom-right table

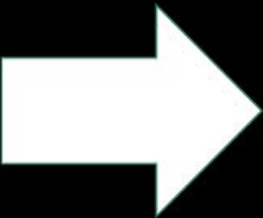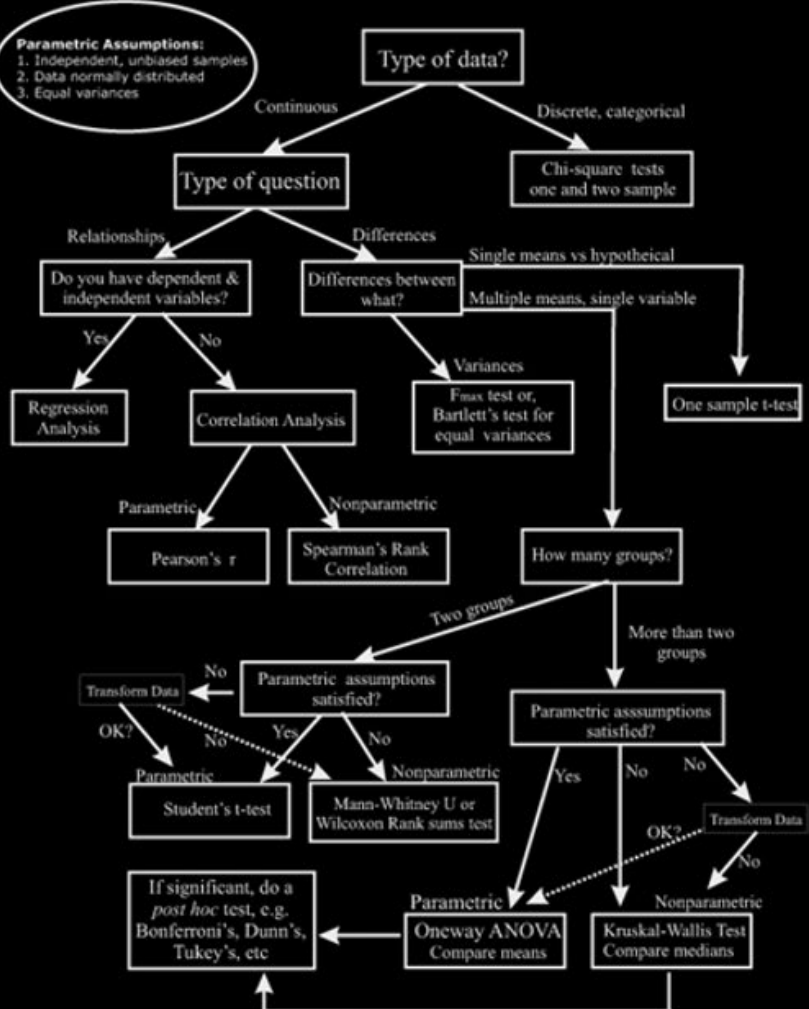|	Drug A			Drug B		
	Heart attack	No heart attack	% of patients with heart attacks	Heart attack	No heart attack	% of patients with heart attacks
Female	1	19	5.0%	3	37	7.5%
Male	12	28	30.0%	8	12	40.0%
Total	13	47	21.7%	11	49	18.3%

Gender

Women less at risk

Women prefer drug B

Type of drug

Controlling for gender, drug A □□ less risk

Heart attack

Next time: GSMS 15 March
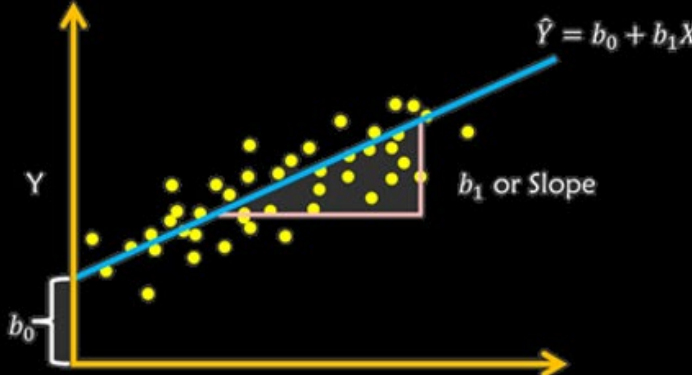
Background knowledge?
Small effects?
Noisy data?

Go Bayes!

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeerch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2013). Labor market returns to early childhood stimulation: A 20-year followup to an experimental intervention in Jamaica. Institute for Research on Labor and Employment working paper #142-13

Gelman, A. (2013). Childhood intervention and earnings. Symposium, 3 Nov. www.symposium-magazine.com/childhood-intervention-and-earnings/

Gelman, A. (2018). The failure of null hypothesis significance testing when studying incremental changes, and what to do about it. *Personality and Social Psychology Bulletin* **44**, 16–23.