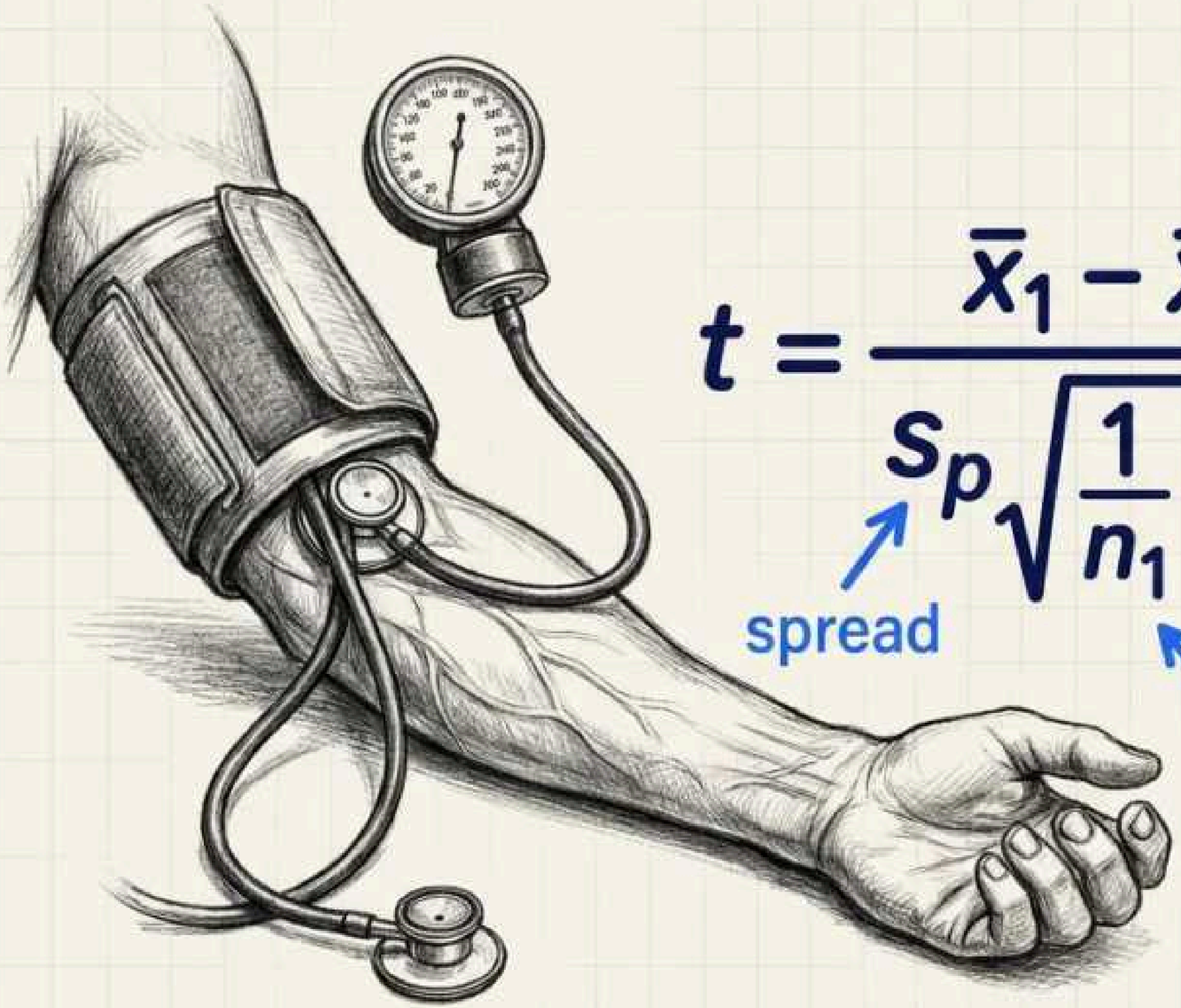
Comparing Numerical Outcomes

¶ 5.1



$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

means

spread

size

The Goal: Comparing treated vs. untreated (or exposed vs. unexposed).

Outcome Type: Numerical (e.g., Blood Pressure).

2 Groups: Use the t-test (accounts for mean, spread, and sample size).

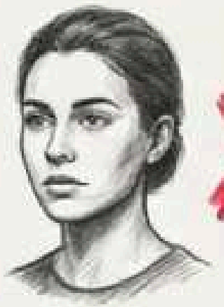
>2 Groups: Use Analysis of Variance (ANOVA).

Gemini Notebook

Comparing Categorical Outcomes



Outcome Type: Categorical
(e.g., Recovered / Not Recovered).



Standard Comparison:
Comparing proportions between groups

Chi-squared test

$$\chi^2 = \frac{\sum (O_i - E_i)^2}{E_i}$$

Very Small Numbers:
When counts are extremely low

Fisher's exact test

Small & Skewed Data:
When data bypasses the normal distribution

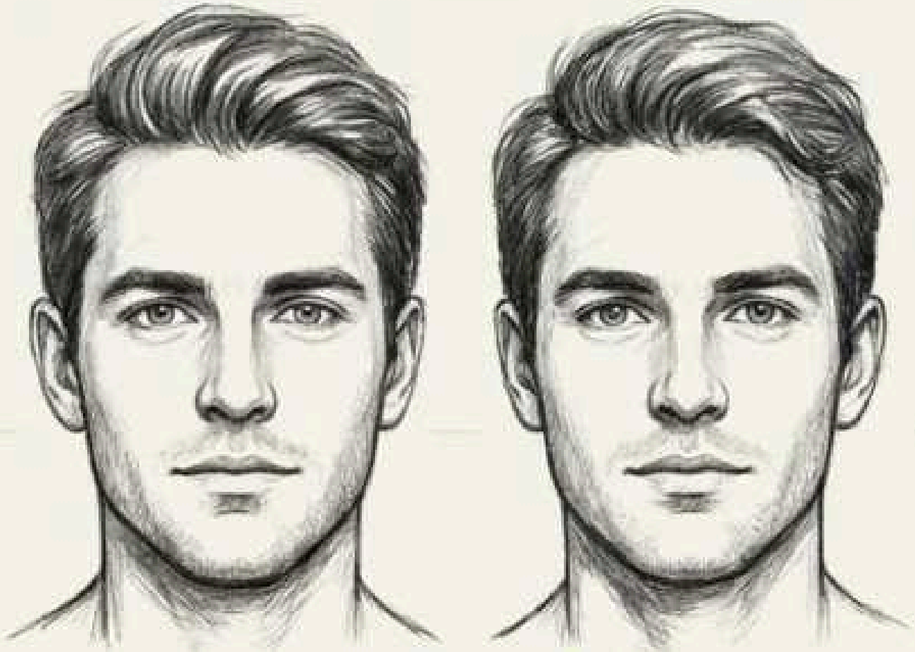
Non-parametric tests

Uses ranks instead of raw values to avoid relying on normal distribution.

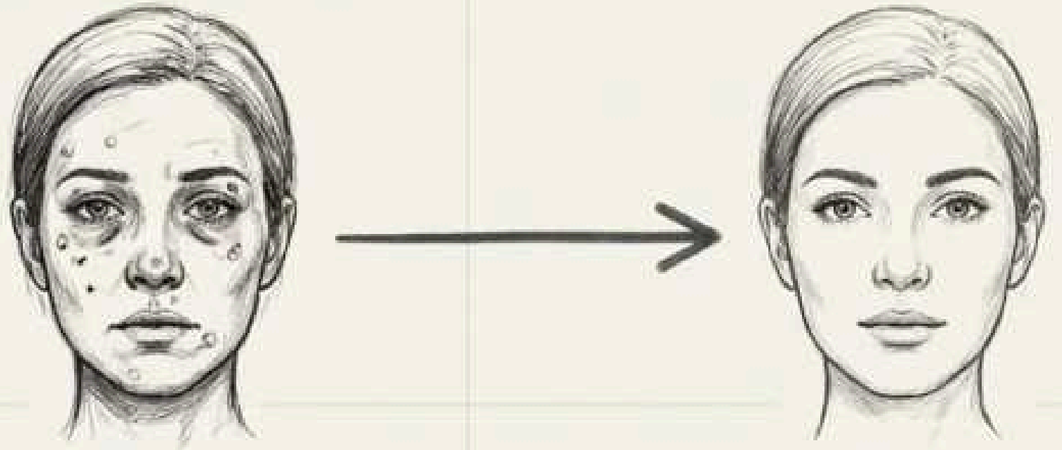
Gemini Notebook

The Power of Paired Data

¶ 5.3



Paired



Before Treatment

After Treatment

$$\Delta = x_{after} - x_{before}$$

Mechanism: Analyzes differences within each pair rather than between broad groups.

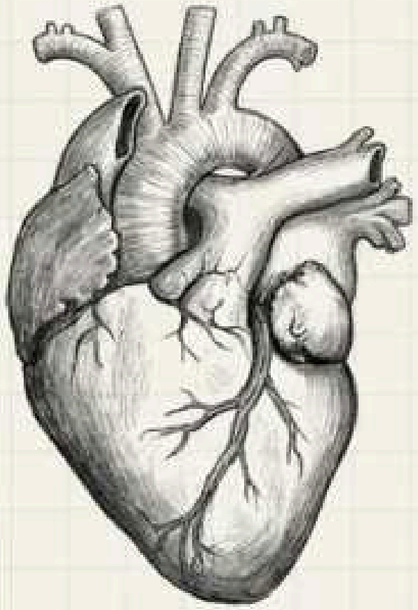
The Advantage: Every patient (or twin) acts as their own control.

The Result: Removes the baseline variation between different people meaning much smaller studies can successfully detect treatment effects.

Gemini Notebook

Defining "Risk" in Medicine

$$\text{Risk} = \frac{\text{People experiencing event}}{\text{Total population over given time}}$$



5 Years

Definition: The proportion of people who experience an event over a given time.

Clinical Scenario: Tracking heart attacks over a 5-year period.

Untreated Group: 10% Risk (10 in 100 patients experience the event).

Treated Group: 8% Risk (8 in 100 patients experience the event).

Note: These two baseline risks form the foundation for all impact measurements.

Gemini Notebook

Relative Risk vs. Absolute Risk Reduction



Relative Risk (RR)

Ratio of the two risks -> 8% / 10% = 0.8

The treatment reduces risk by 20% in relative terms.

Impression: Sounds massive and highly effective.

Absolute Risk Reduction (ARR)

Difference between the risks -> 10% - 8% = 2%

The treatment reduces risk by exactly 2 percentage points.

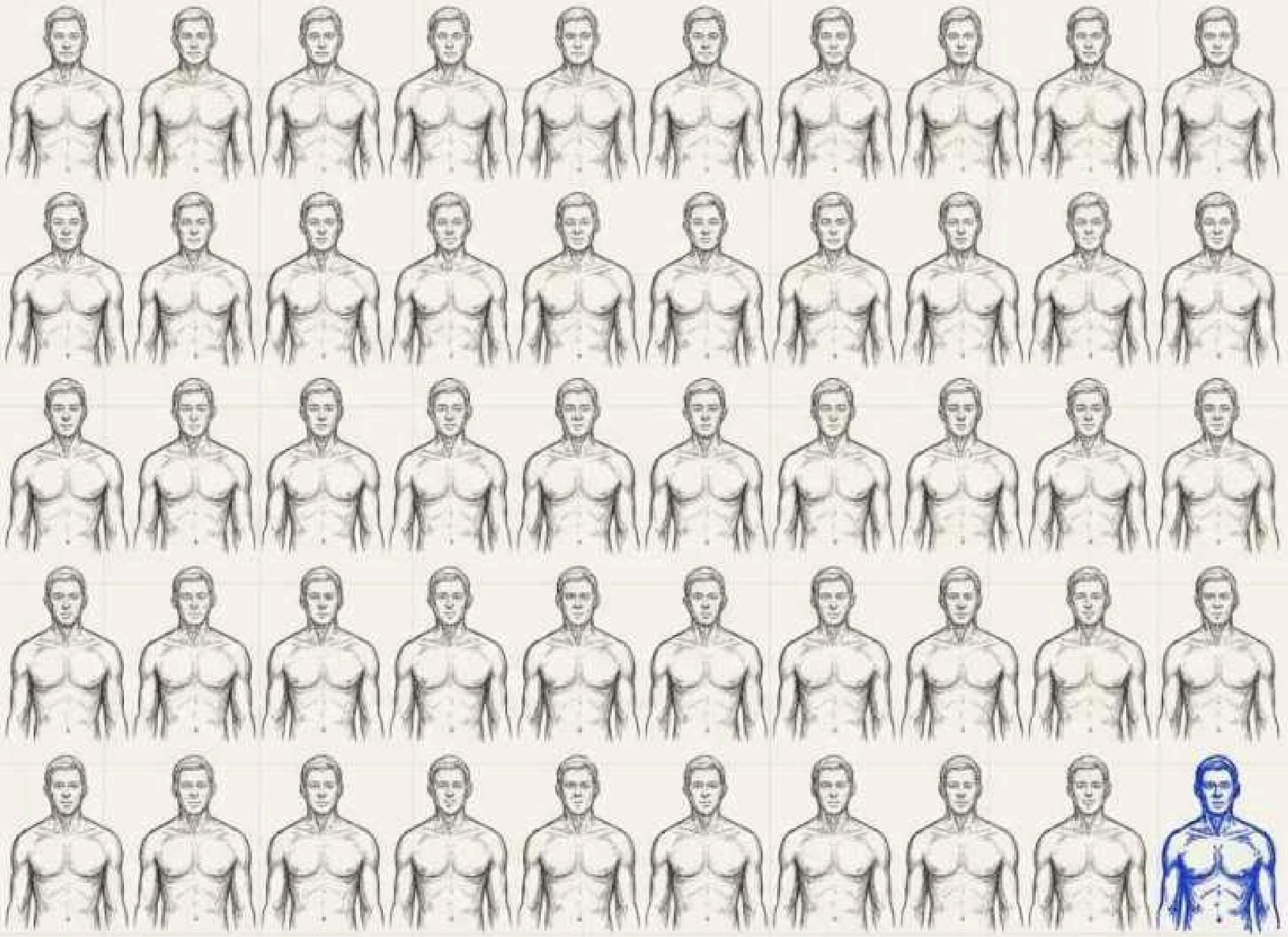
Impression: Sounds modest and incremental.

Both are mathematically correct; both create very different clinical impressions.

Gemini Notebook

Number Needed to Treat (NNT)

¶ 5.6



$$NNT = \frac{1}{ARR}$$

From previous example:

$$\frac{1}{0.02} = 50$$

The Meaning: Turns absolute risk into a practical figure. 50 patients must take the treatment for 5 years to prevent one single heart attack.

The Baseline Caveat

If the baseline risk was only 1%, that same 20% relative reduction yields an ARR of 0.002.

New Calculation:
NNT = 1 / 0.002 = 500.
(Now, 500 must be treated to save 1).

Gemini Notebook

Why Absolute Numbers Matter



Favors Relative Risk for drama



The Illusion:
"Halving the risk" of a rare disease changes an individual's actual chances very little.

The Clinical Trade-off:
If you only look at relative risk, you may expose 100% of treated patients to real drug side effects for a negligible absolute benefit.

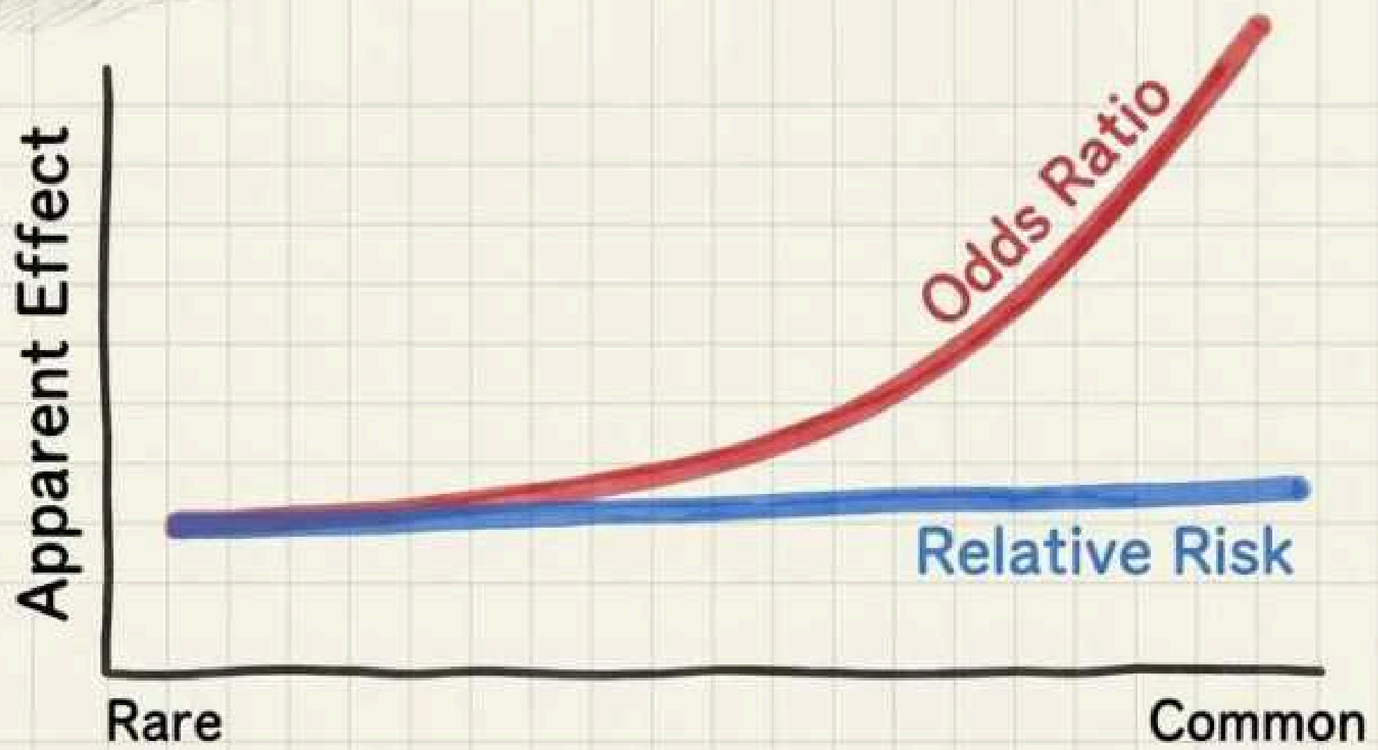
Core Rule:
Doctors and patients should always ask for absolute numbers.

Gemini Notebook

Understanding Odds Ratios



$$\text{Odds} = \frac{\text{Chance it happens}}{\text{Chance it does not}}$$



Usage: Often reported in Case-Control Studies.

Rare Outcomes:
Odds Ratio ≈ Relative Risk.

Common Outcomes:
Odds Ratio mathematically exaggerates the apparent effect.

Takeaway: Must be interpreted with care; it is not always interchangeable with Relative Risk.

Gemini Notebook

Correlation is Not Causation

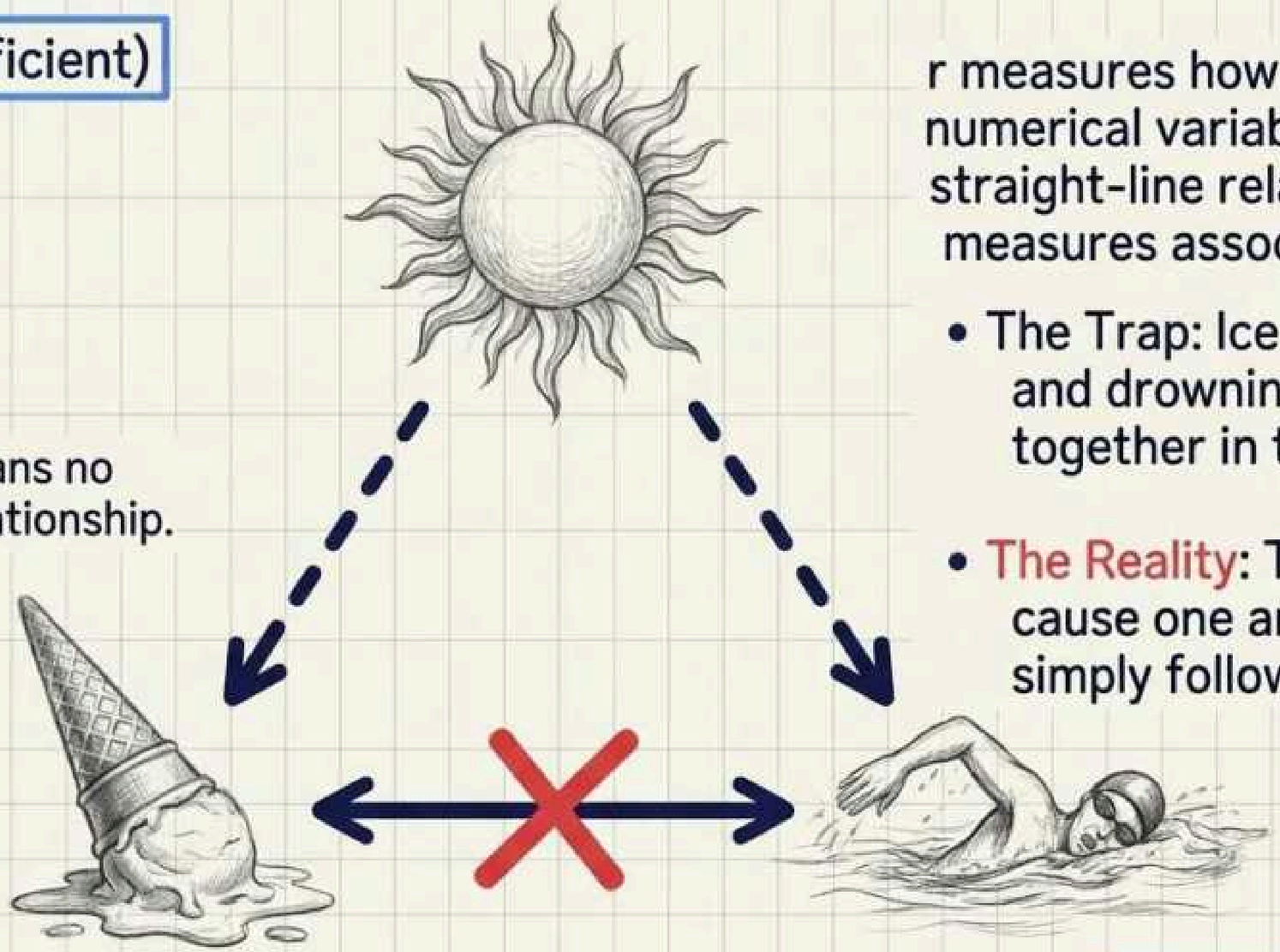
¶ 5.9

r (Correlation Coefficient)

+1

0 ← r ≈ 0 means no linear relationship.

-1

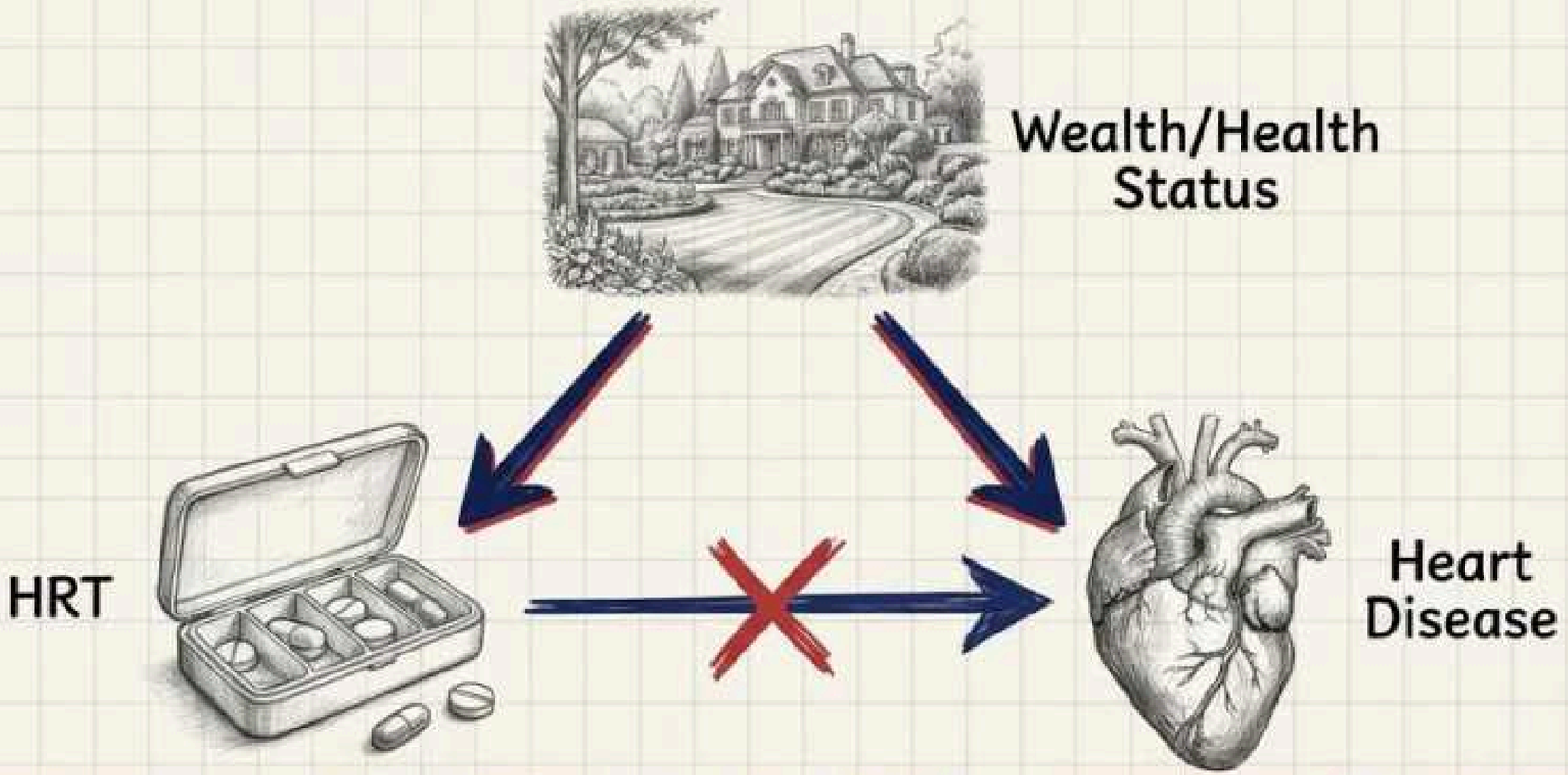


r measures how closely two numerical variables follow a straight-line relationship. It measures association only.

- The Trap: Ice cream sales and drownings rise together in the summer.
- The Reality: They do not cause one another; both simply follow hot weather.

Gemini Notebook

The Threat of Confounding



Definition: A confounder is a third factor linked to both the exposure and the outcome, creating a false association.

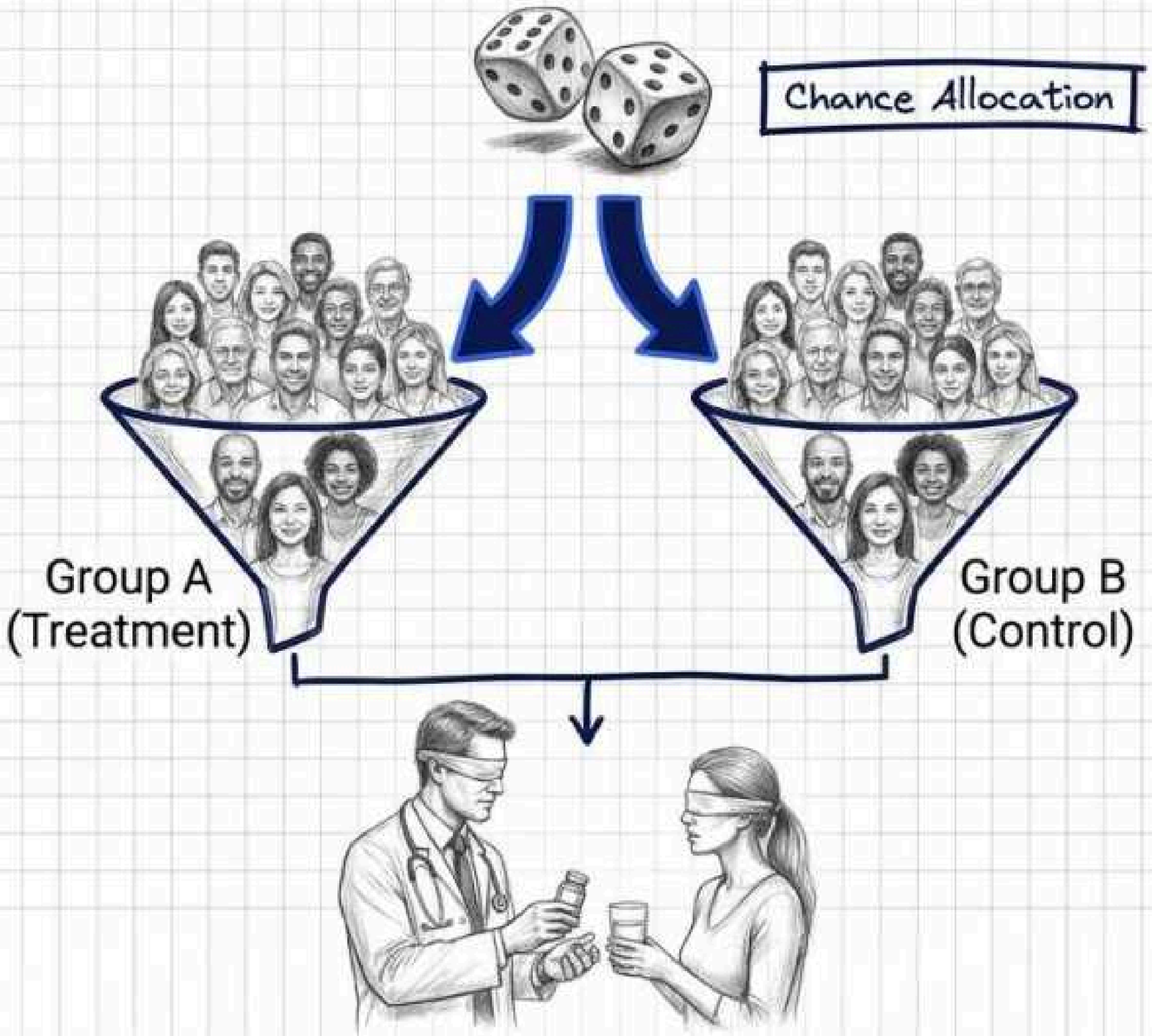
Historical Error: Early observational studies suggested HRT protected women against heart disease.

The Confounder: Women taking HRT tended to be healthier and wealthier at baseline.

The Truth: Large trials later proved HRT offered no such protection.

Gemini Notebook

Randomised Controlled Trials (RCTs)



The Defense:

Randomisation is the strongest defense against confounding.

How it Works:

By allocating treatment by chance, groups become similar in every respect—both known and unknown—apart from the treatment itself.

Blinding:

Neither patients nor doctors know who received which treatment. This prevents psychological expectations from distorting the results.

Gemini Notebook

Lecture Synthesis & What's Next

1. Compare Groups Accurately:
Choosing the right test is driven entirely by **data type** (numerical vs. categorical).

2. Describe Effects Fully:
Always require **Absolute measures** (ARR, NNT) alongside **Relative measures** (RR, OR) for the complete picture.

3. Prove Causation:
Association ≠ Causation.
Randominisation is the surest auotay to separate them and eliminate confounding.



Gemini Notebook