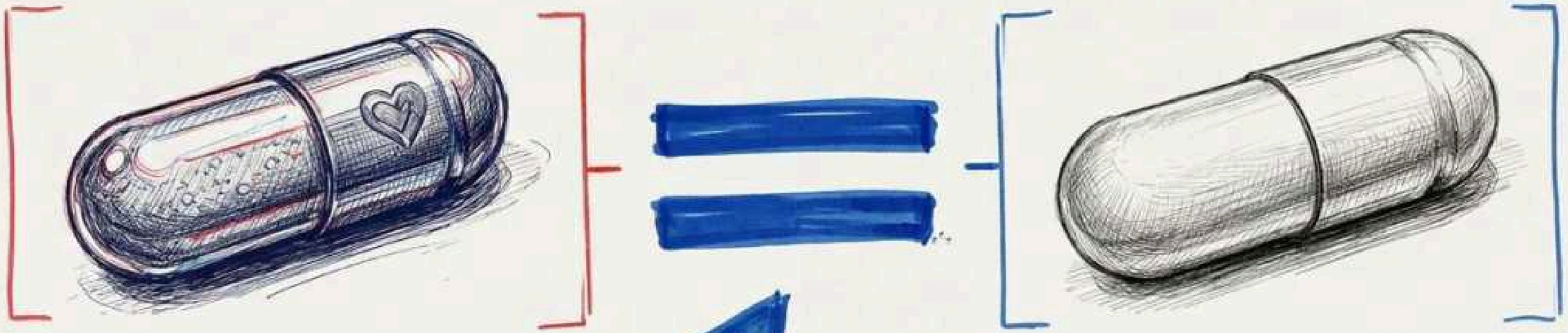
# The Null Hypothesis ($H_0$)

¶ 4.1



**The Premise:**
A hypothesis test asks if data is consistent with a specific claim.

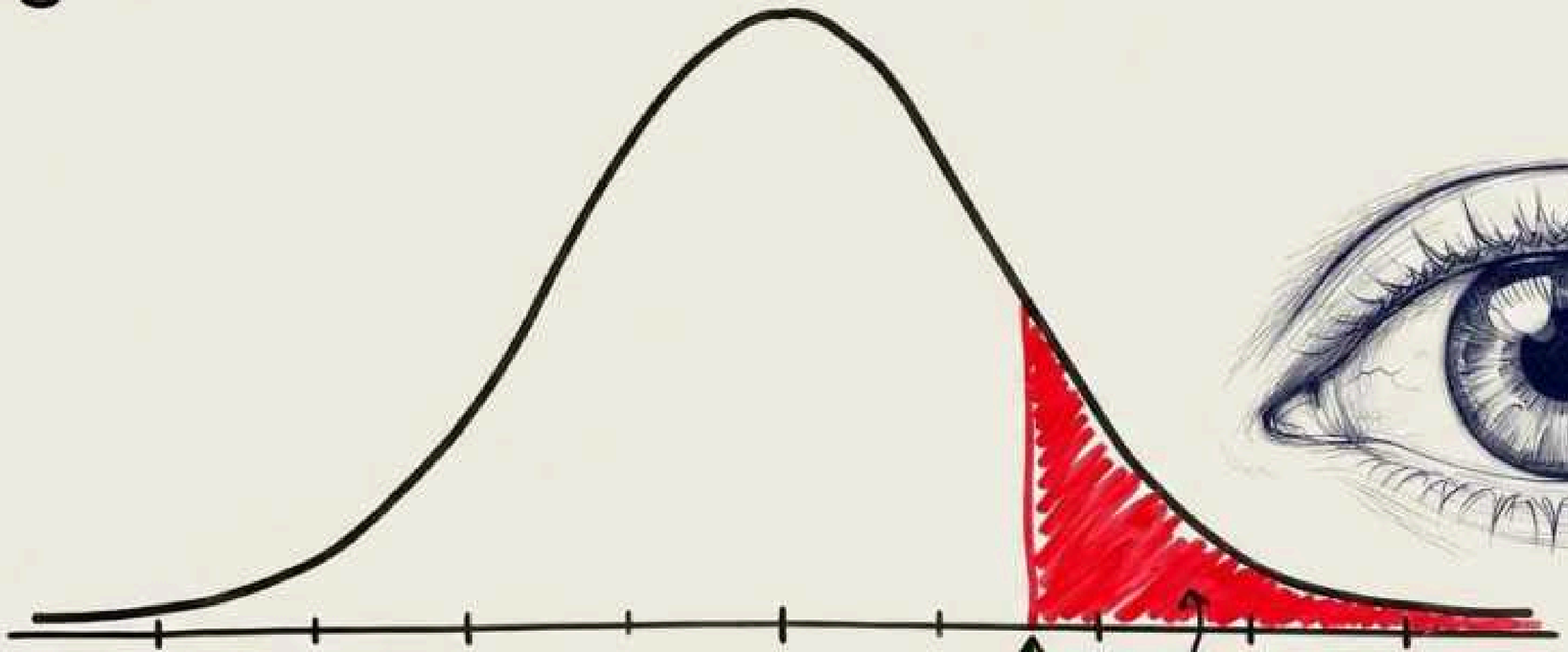
**The Null ($H_0$):**
The assumption that there is no effect or difference.

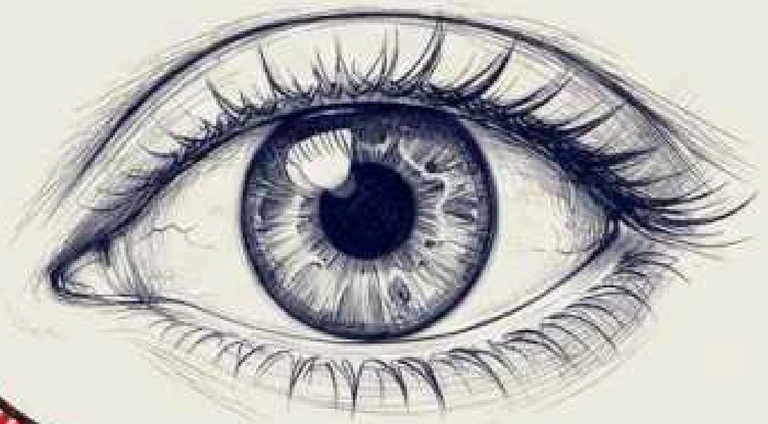
In a clinical trial, the null hypothesis assumes the new drug makes no difference to the outcome compared to a placebo.

Gemini Notebook

# Defining the P-Value

¶ 4.2



$H_0$: Null Hypothesis

**The True Definition:**
The probability of seeing a result at least as extreme as the one observed, assuming the null hypothesis ($H_0$) is true.

**The Interpretation:**
A small p-value means the observed data would be highly surprising if there were really no effect.

↳ Highly unlikely event under $H_0$
Strong evidence against $H_0$

Area = p-value

**The Arbitrary Line:** ξ?
By long convention, $p < 0.05$ is designated as "statistically significant," though this cut-off is entirely arbitrary.

↖ Historical precedent, not a hard rule!

↖ Usually, this area is < 5%

Gemini Notebook

# What a P-Value is NOT

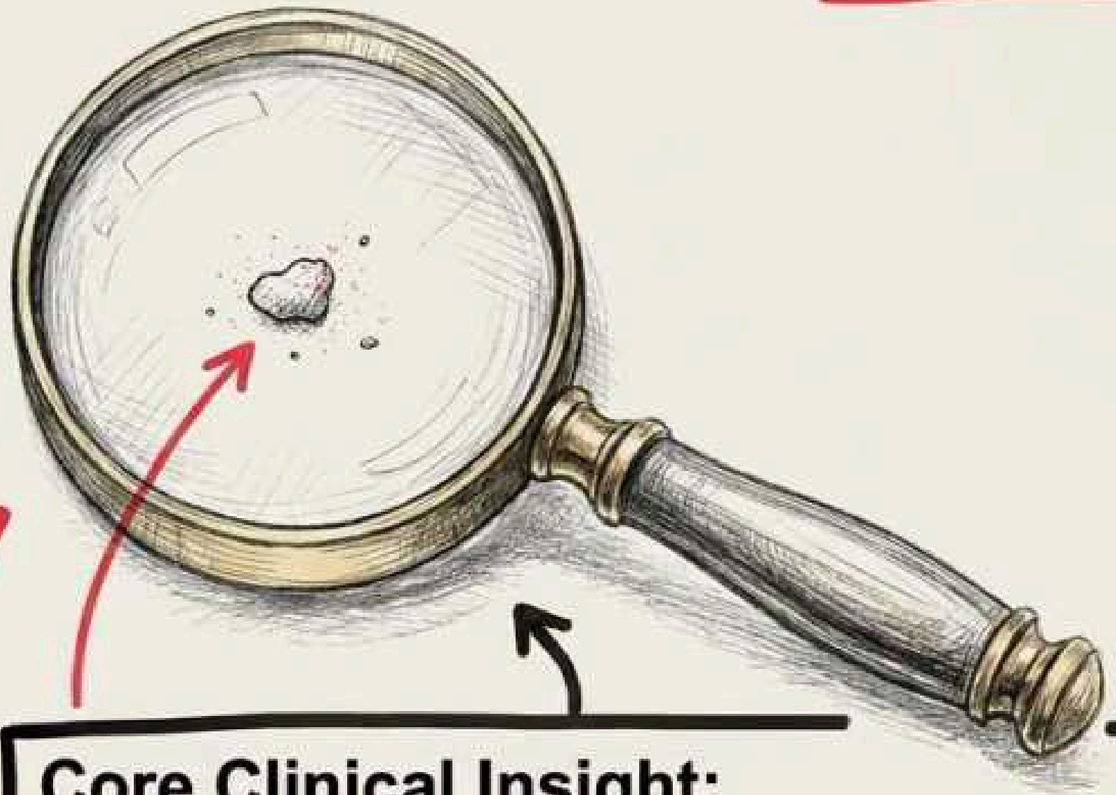
¶ 4.3

## Myth vs Reality

It is NOT the probability that the null hypothesis is true.

It is NOT the probability that the result happened by chance.

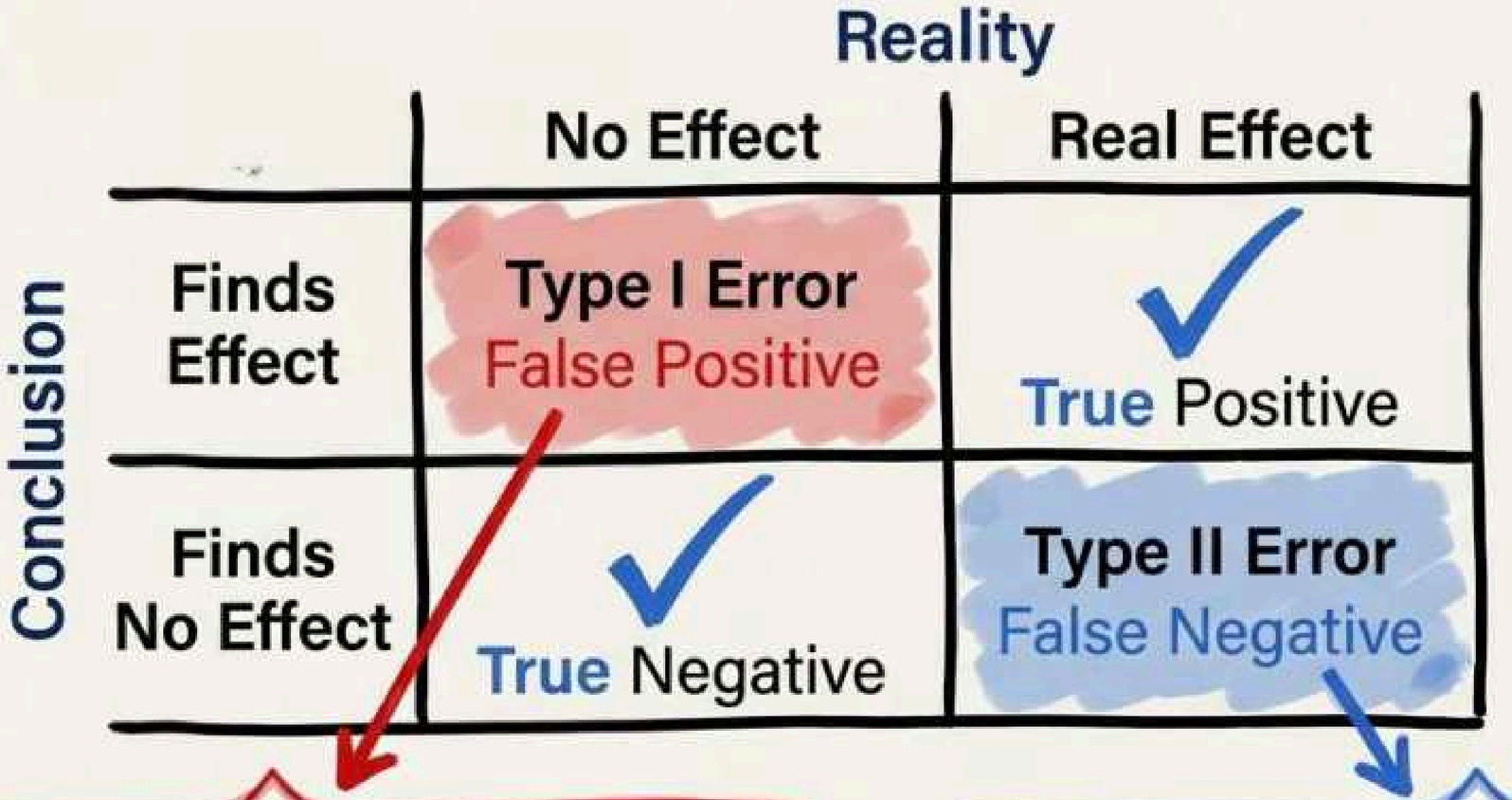
It is NOT a measure of the effect's size or clinical importance.



**Core Clinical Insight:**
A tiny, clinically worthless effect can can still yield a very small p-value if study's sample size is large enough.

# The Two Errors in Testing

¶ 4.4



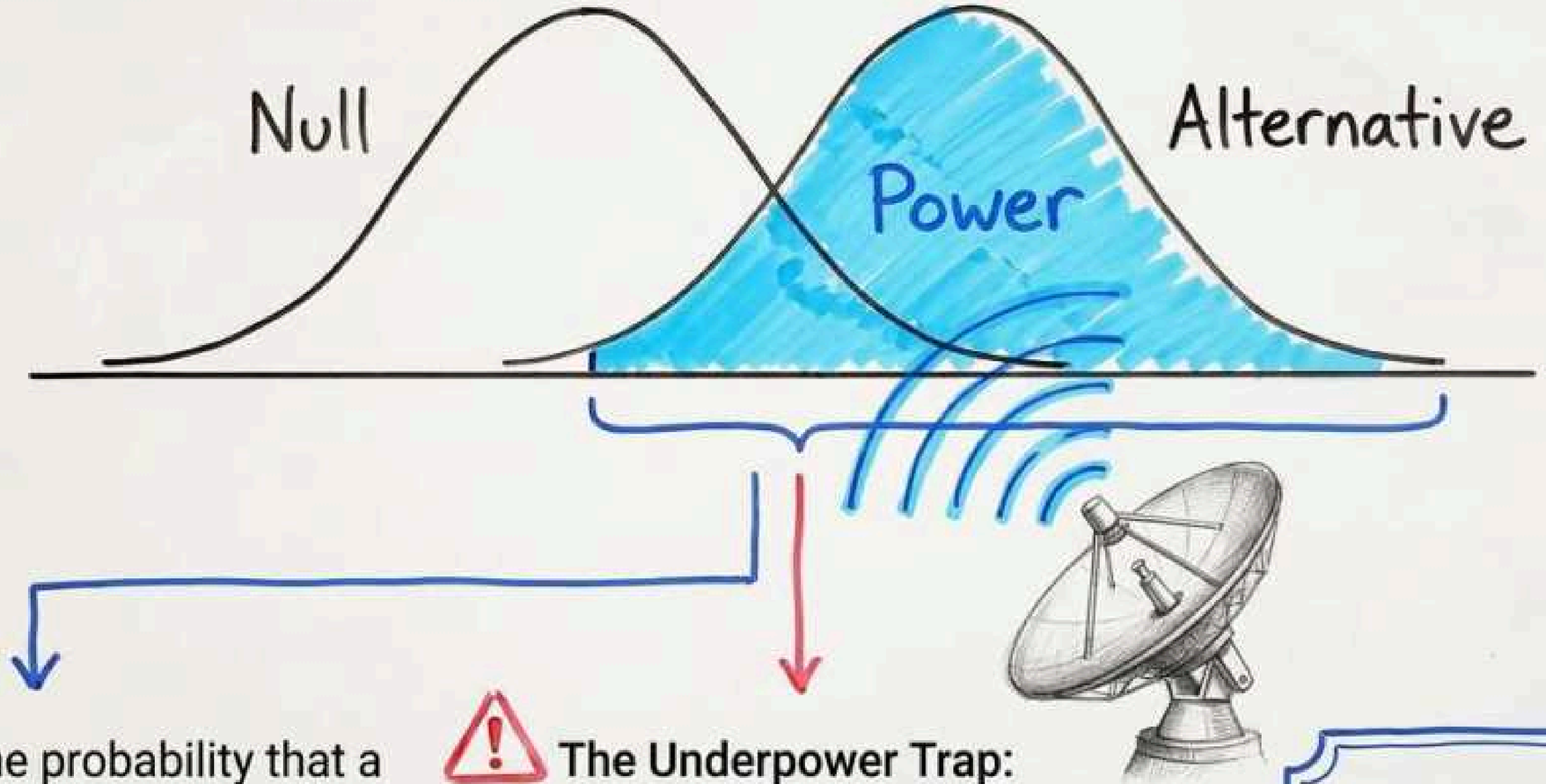

**Type I Error (False Positive)**: Concluding there is an effect when there is none. The risk is set by the significance level (usually 5%).

**Type II Error (False Negative)**: Missing a real effect. The risk depends on three factors: the size of the effect, the variability of the data, and the sample size.

Gemini Notebook

# Statistical Power (1 - β)

¶ 4.5




**Definition:** The probability that a study will detect a real effect of a given size.
Studies are usually designed with 80% or 90% power.

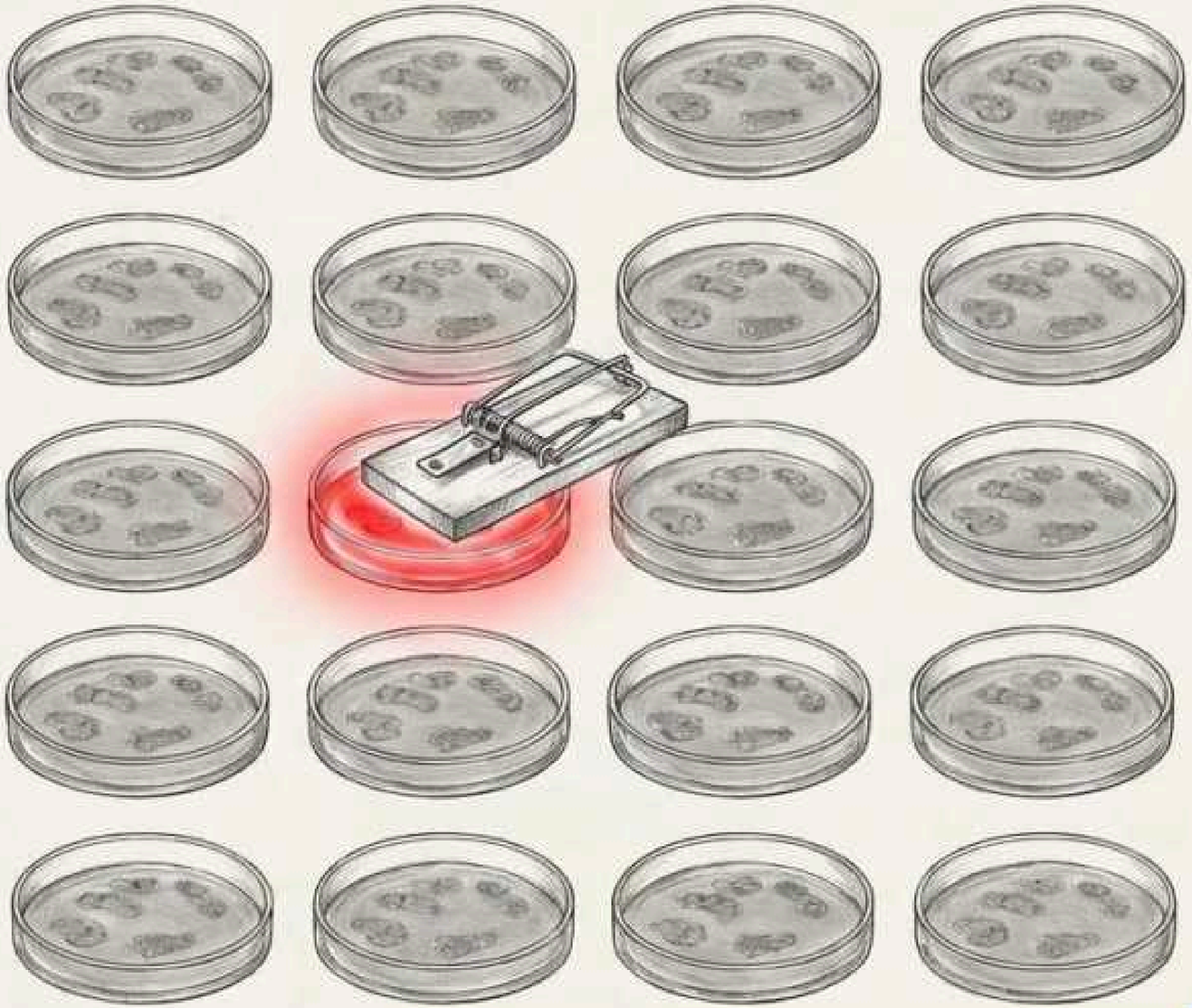
**The Underpower Trap:**
An underpowered study that finds no significant difference has not shown that a treatment doesn't work; it has simply failed to find out.

**The Golden Rule:**
Absence of evidence is not evidence of absence.

Gemini Notebook

# The Multiple Testing Trap

¶ 4.6



**The Math of Chance:** If 20 independent comparisons are made where no real effects exist, we expect about 1 to be **significant at the 5%** level by pure chance alone.

**The Trap:**
Studies that test many outcomes or subgroups and report only the significant ones produce impressive-looking, but entirely false, findings.

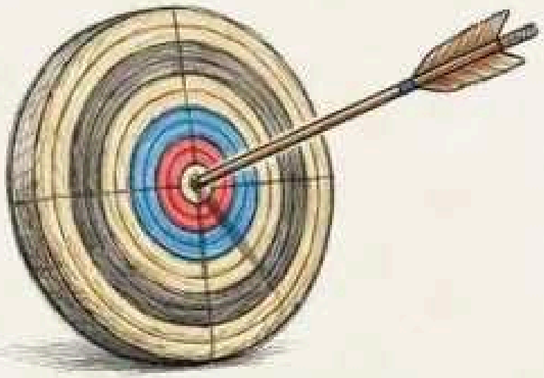
Gemini Notebook

# Safeguarding the Truth

¶ 4.7



Name a **single primary outcome** before the study starts and publicly register plans.



When **multiple comparisons** are unavoidable, **use stricter thresholds** (like the **Bonferroni** correction, dividing the significance level by the number of tests).

Bonferroni: $\alpha_{new} = \frac{0.05}{n}$

Treat all unplanned findings strictly as ideas to test, not as verified conclusions.

Gemini Notebook

# Statistical vs. Clinical Significance

¶ 4.8

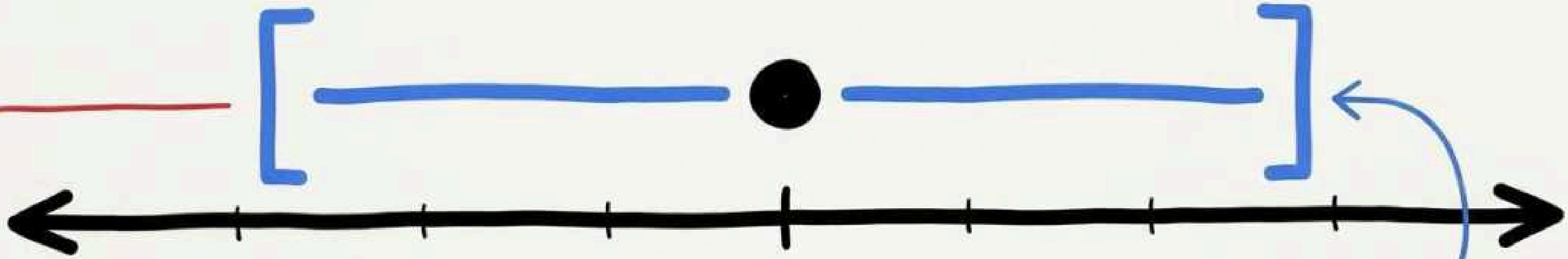



- Statistical and clinical significance are fundamentally different questions.
- A 1-point blood pressure drop in 50,000 people may be statistically significant, yet matter little to an individual patient.
- **The Doctor's Mandate: Always ask how large the effect is, and if it justifies the costs and side effects.**

Gemini Notebook

# The Power of Confidence Intervals

¶ 4.9



**Why they beat p-values:** An interval shows both whether an effect is likely to be real AND how big it might plausibly be.

Provides the actual **range** of the effect's magnitude.

Many medical journals now mandate reporting effect sizes with confidence intervals rather than relying on p-values alone.

Gemini Notebook

# The Bayesian Alternative

¶4.10



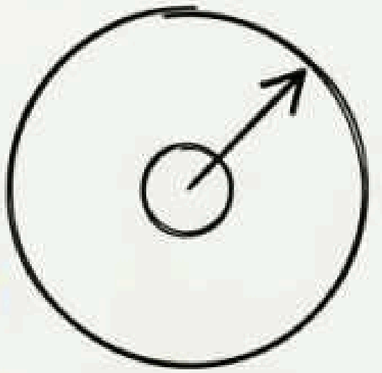
Prior Explicit
Starting Belief

+



New Data

=



Direct Probability
of Effect Sizes

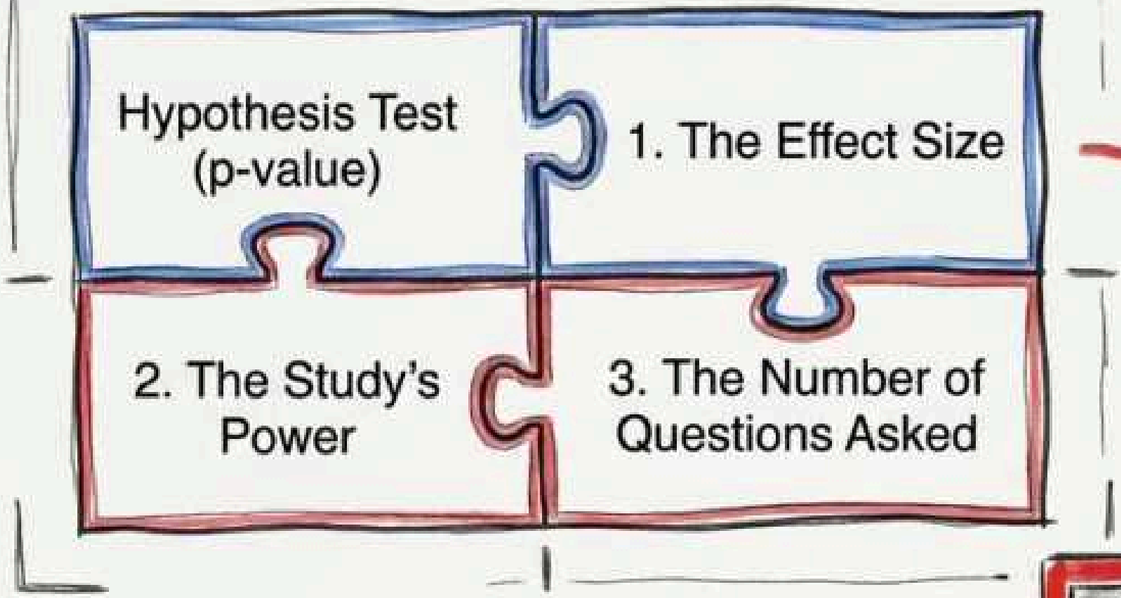
- Combines prior knowledge with new data to output direct probabilities.
- Often yields answers much closer to what doctors and patients actually want to know.
- **The Trade-off**: Requires the statistician to acknowledge and quantify an explicit starting belief.

Gemini Notebook

# The Clinical Statistician's Toolkit

¶ 4.11

A hypothesis test is just one tool for judging surprise under a no-effect assumption. It must never be read alone.




**Preview — Lecture 5:**
**Applying these exact tools to compare groups and measure how strongly clinical variables are linked.**

Gemini Notebook