P-Hacking and Other Statistical Malfeasance

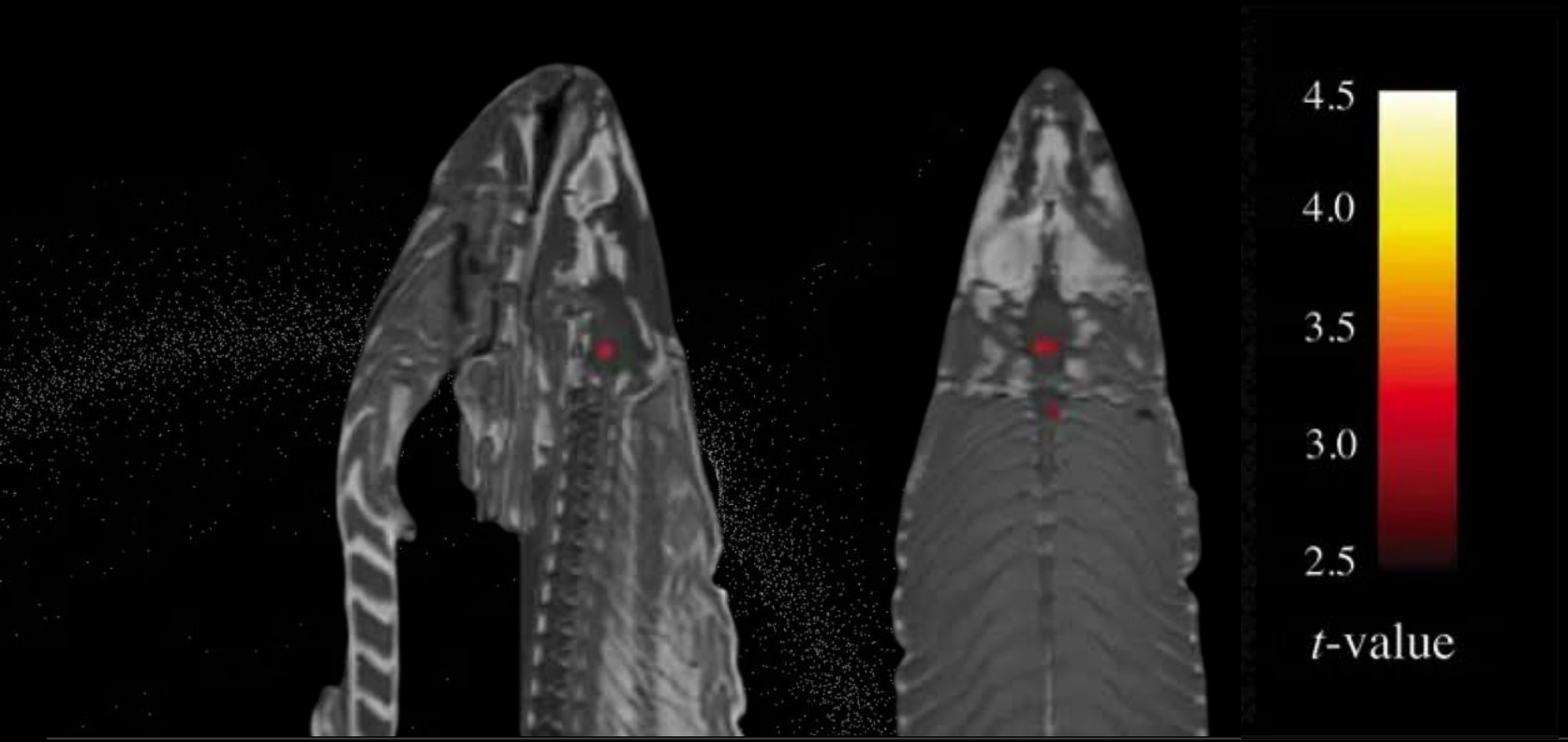
A highly significant talk (≥5σ, after a few reasonable cuts)

Dillon Bass, ASTR 8500

University of Virginia

Apr, 2026

"Significant" Findings



Key Ideas

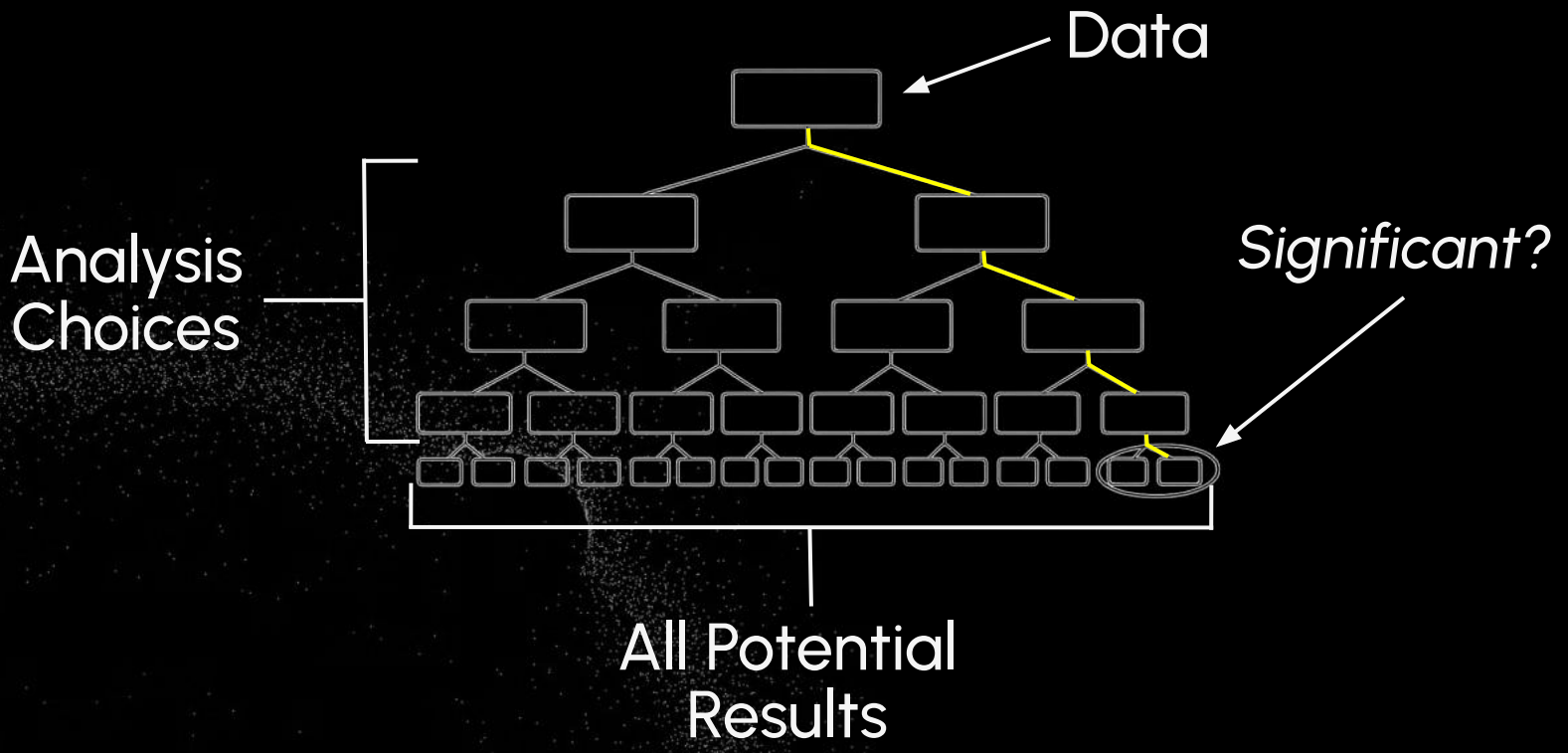
1. Statistical malfeasance is not necessarily malicious, and is extremely common

2. Statistical significance is not just a property of the data—it is a property of the analysis process.

What is P-Hacking?

"...performing many statistical tests on the data and only reporting those that come back with significant results."

- trying multiple cuts
- different models
- stop when significant
- selective reporting


Data
Analysis Choices
Significant?
All Potential Results

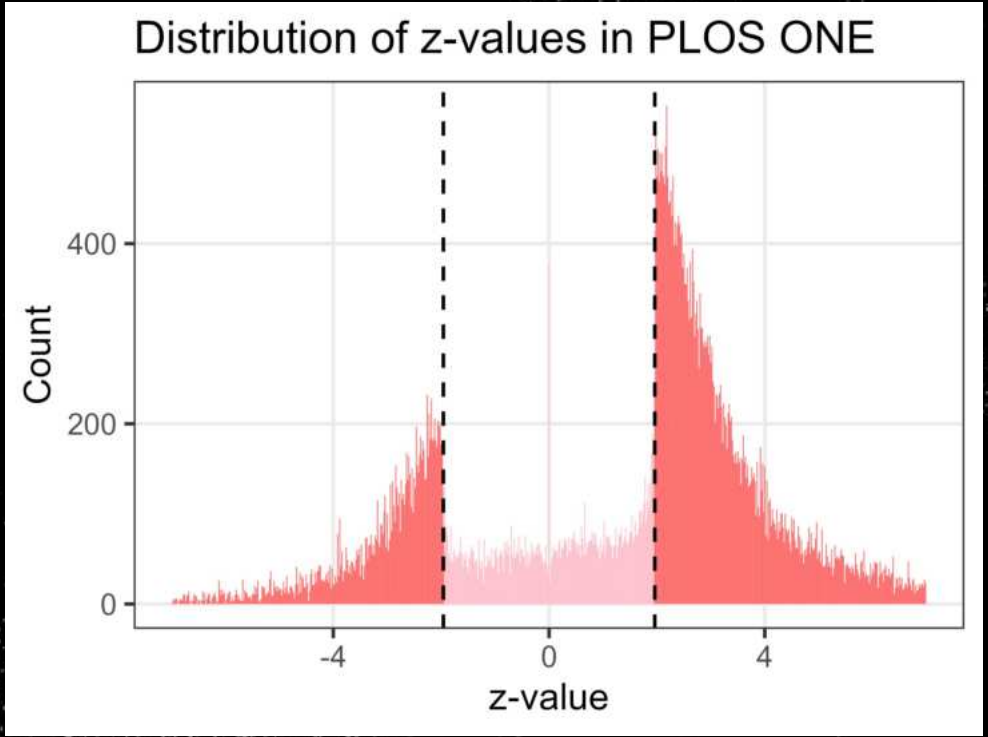

If you run enough analyses...
IT MEANS THAT WHATEVER CAN HAPPEN WILL HAPPEN.

Other Common Sources

- Drawing conclusions from data
  - Outlier Removal
  - Optional stopping
  - Post-hoc grouping
  - More...

See how these might be insidious?? They don't even sound like bad things...

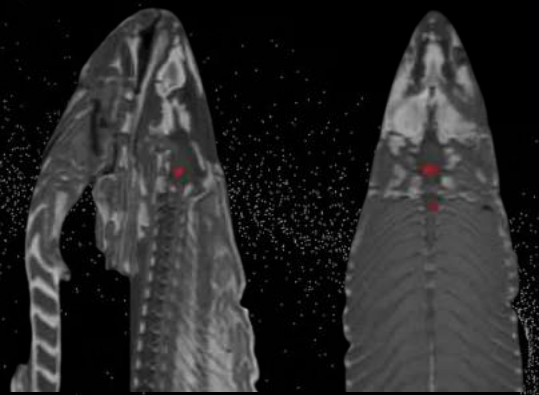
On a Large Scale...



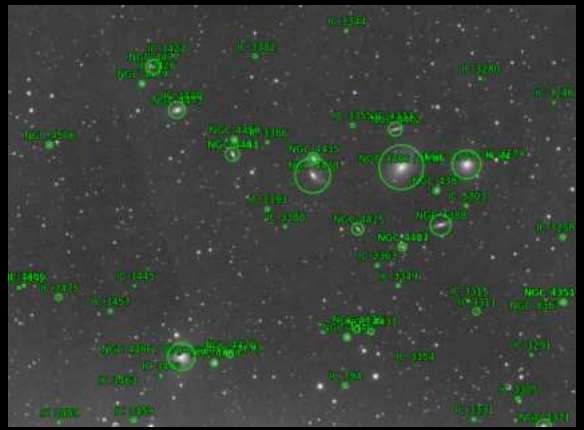
In Astro...

- We search over sky regions, frequencies, and big datasets
- We try multiple pipelines, models, and systematics corrections
- We tune cuts, thresholds, and cleaning choices
- We choose what to report after seeing the data

So many ‘tests’…



MRI Voxels



Sky Samples

Mitigation Strategies

1. Be aware of hidden trials

(Significance depends on how many analyses you could have run)

2. Make analysis choices explicit and fixed

(Predefine cuts, models, and pipelines before looking at results)

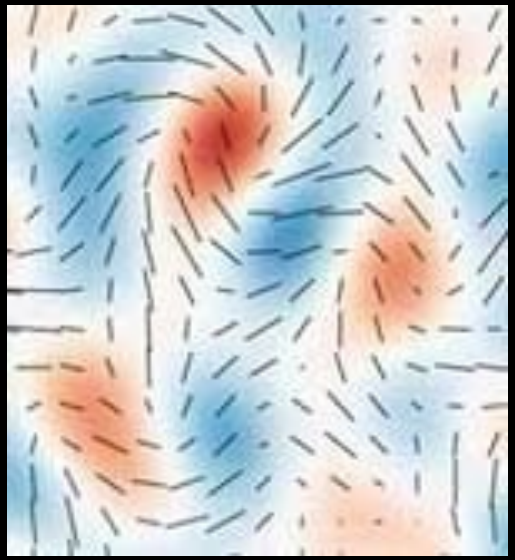
3. Test robustness, not just significance

(Does the result survive reasonable alternative choices?)

In the end...

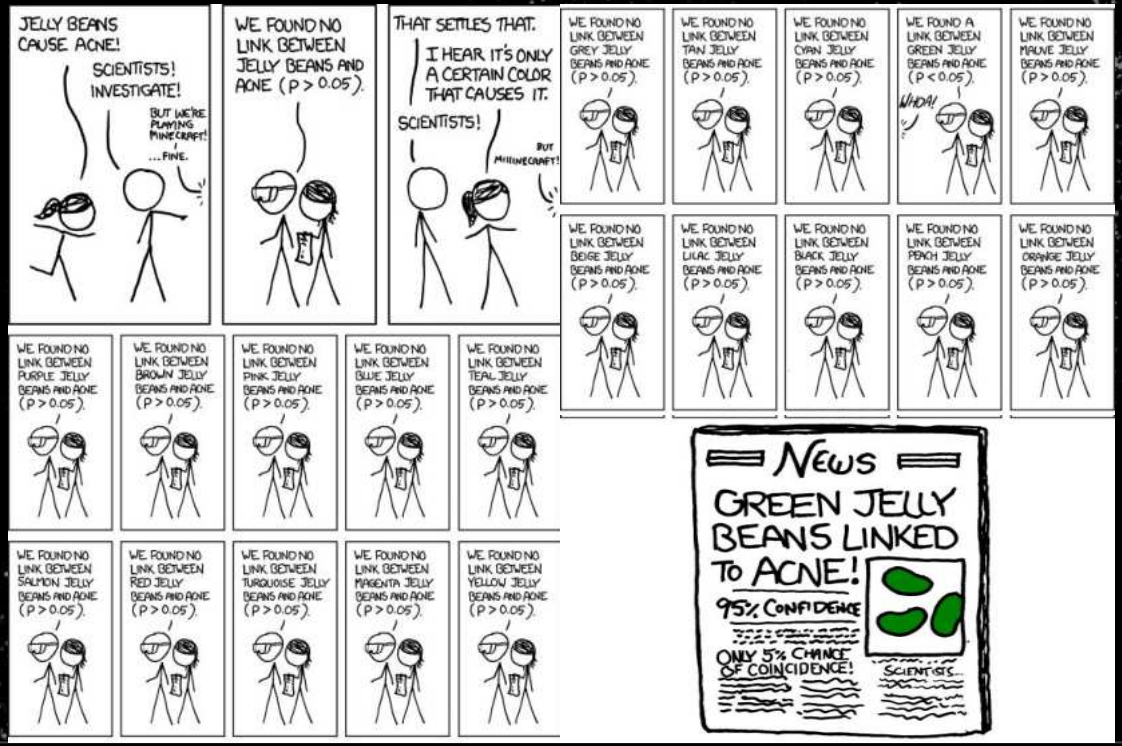
We often think that our integrity as researchers is enough to avoid biased results...

But the tools that we use to do research every day can also be statistically dangerous if we aren't careful.



5σ BICEP2 'detection' of primordial gravitational waves

Questions?


JELLY BEANS CAUSE ACNE!
SCIENTISTS! INVESTIGATE!
BUT WE'RE PLAYING MINECRAFT!
...FINE.
WE FOUND NO LINK BETWEEN JELLY BEANS AND ACNE (P > 0.05).
THAT SETTLES THAT.
I HEAR IT'S ONLY A CERTAIN COLOR THAT CAUSES IT.
SCIENTISTS!
BUT MINECRAFT!
WE FOUND NO LINK BETWEEN GREY JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TAN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN CYAN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND A LINK BETWEEN GREEN JELLY BEANS AND ACNE (P < 0.05).
WHOA!
WE FOUND NO LINK BETWEEN MAUVE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BEIGE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN LILAC JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLACK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PEACH JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN ORANGE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PURPLE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BROWN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PINK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLUE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TEAL JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN SALMON JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN RED JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TURQUOISE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN MAGENTA JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN YELLOW JELLY BEANS AND ACNE (P > 0.05).
NEWS
GREEN JELLY BEANS LINKED TO ACNE!
95% CONFIDENCE
ONLY 5% CHANCE OF COINCIDENCE!
SCIENTISTS

Sources

- https://www.sciencedirect.com/science/article/pii/S1053811909712029
- https://www.wired.com/2009/09/fmrisalmon/
- https://en.wikipedia.org/wiki/Data_dredging
- https://medianwatch.netlify.app/post/z_values/
- https://www.nature.com/articles/nature.2015.16830
- https://pmc.ncbi.nlm.nih.gov/articles/PMC4359000/
- https://www.nature.com/articles/d41586-025-01246-1
- https://www.science.org/content/article/curtain-falls-controversial-big-bang-result
- https://xkcd.com/882/