

Ethics in science: Experimental design and bias

Frédéric Schütz

Frederic.Schutz@unil.ch

Biostatistics platform, FBM
University of Lausanne

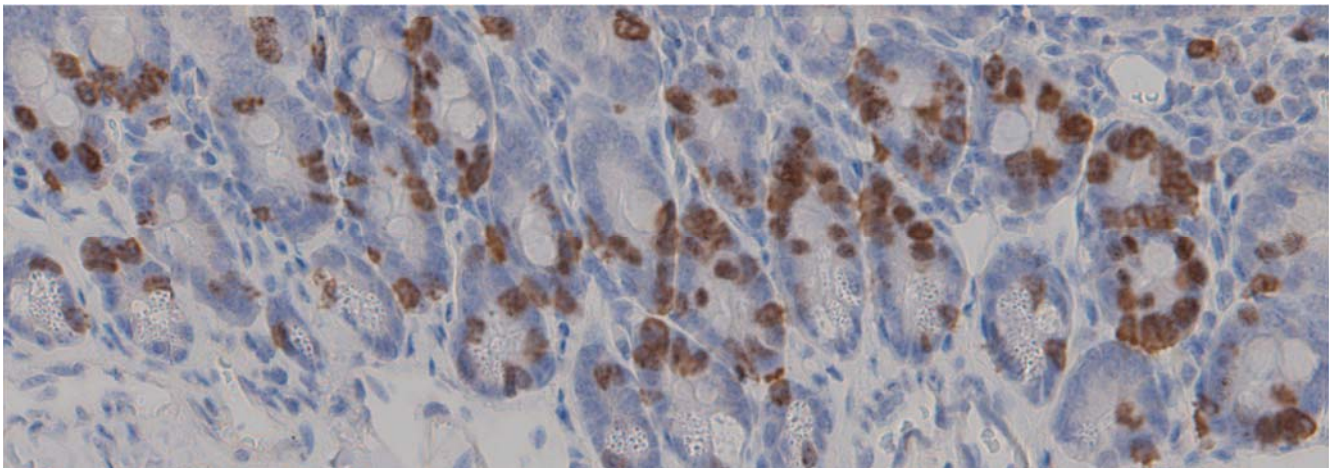


Once upon a time...

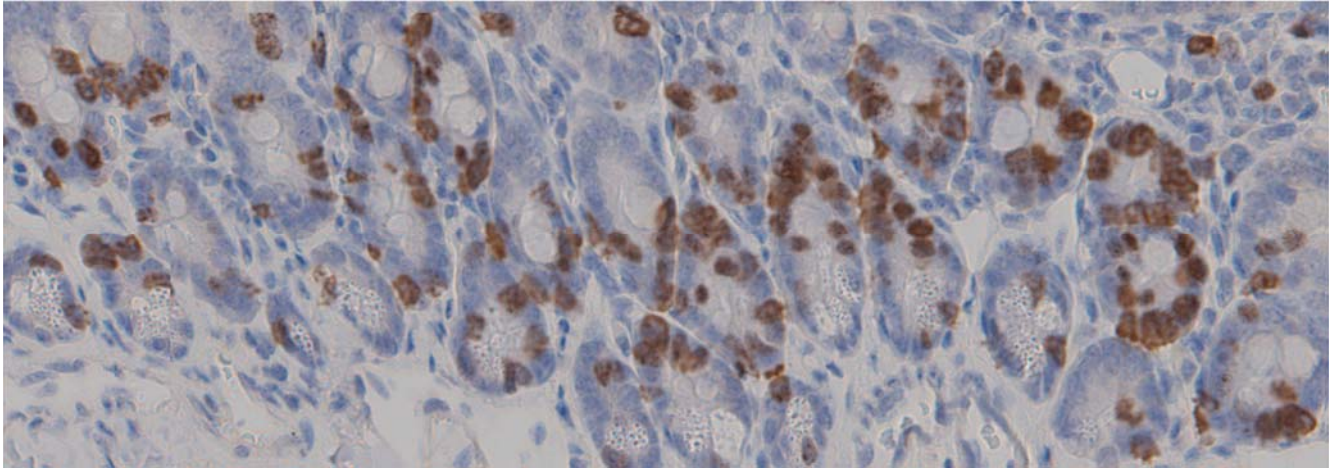
**A class of students were asked
to count the number
of brown nuclei
in two images.**

Image 1

This image was labelled as coming from a **healthy patient**.
We expect to observe a **low** number of brown nuclei.

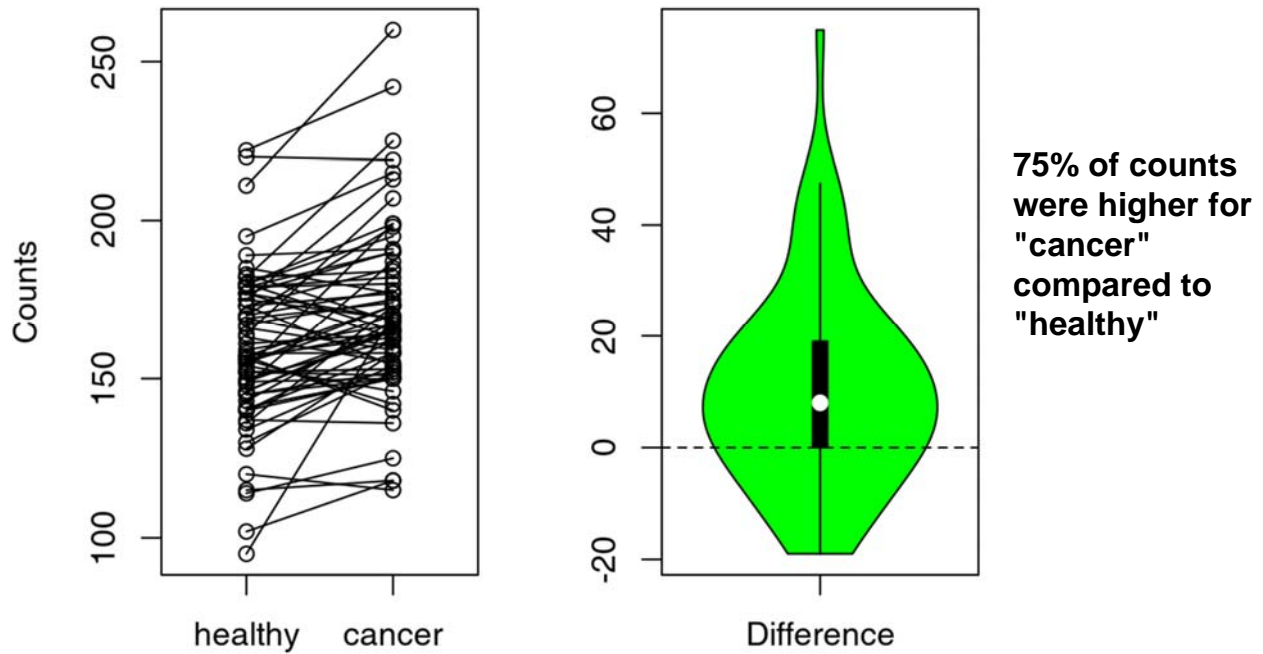


This image was labelled as coming from a **cancer patient**.
We expect to observe a **large** number of brown nuclei.



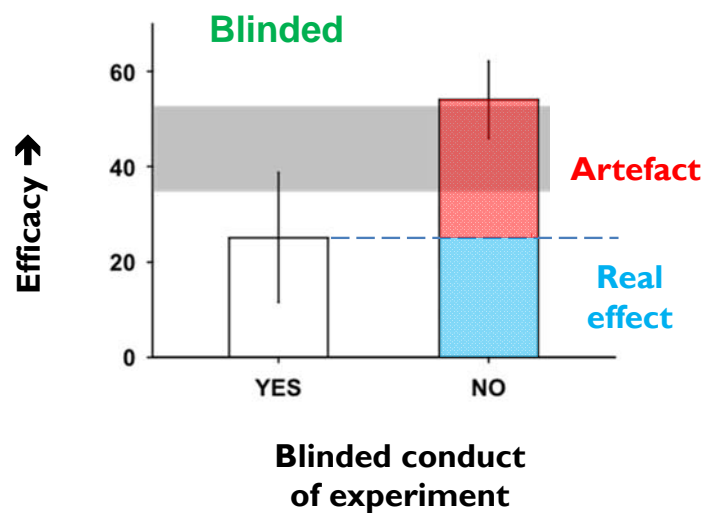
Goal:
Assess whether the students were
(unconsciously) influenced by the
stated expectations

Paired data

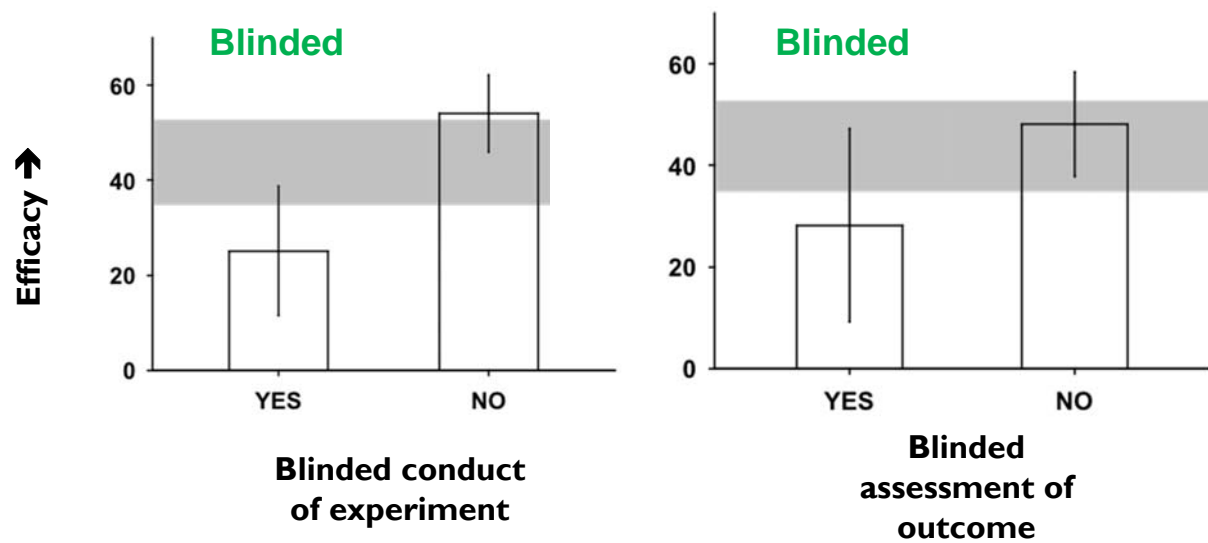


Mean difference: **10.5 units**; p-value: **5×10^{-6}**
(paired Student's t-test; $t = 4.9585$, $df = 68$)

Effect of bias in an animal experiment: the example of stroke



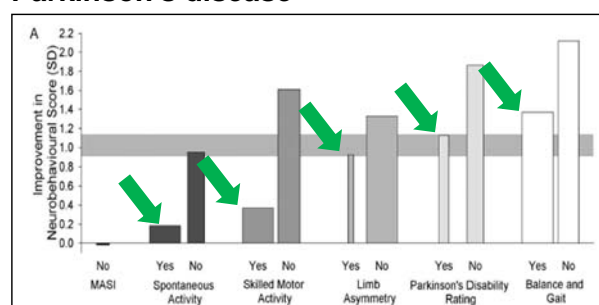
Effect of bias in an animal experiment: the example of stroke



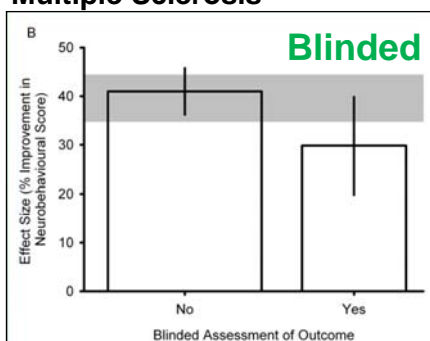
Malcolm R. Macleod et al. Stroke (39), 2009. 2824–2829

Evidence from other domains

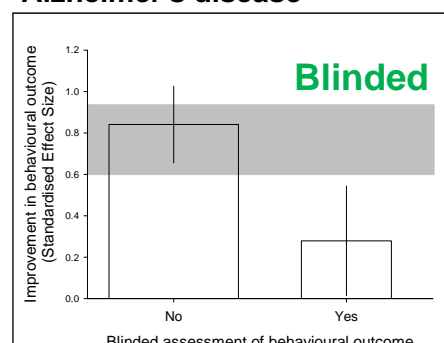
Parkinson's disease



Multiple Sclerosis



Alzheimer's disease



From Malcolm R. Macleod

What about the data analysis step ?

"At least, this part should be unbiased..."

Well...

Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty

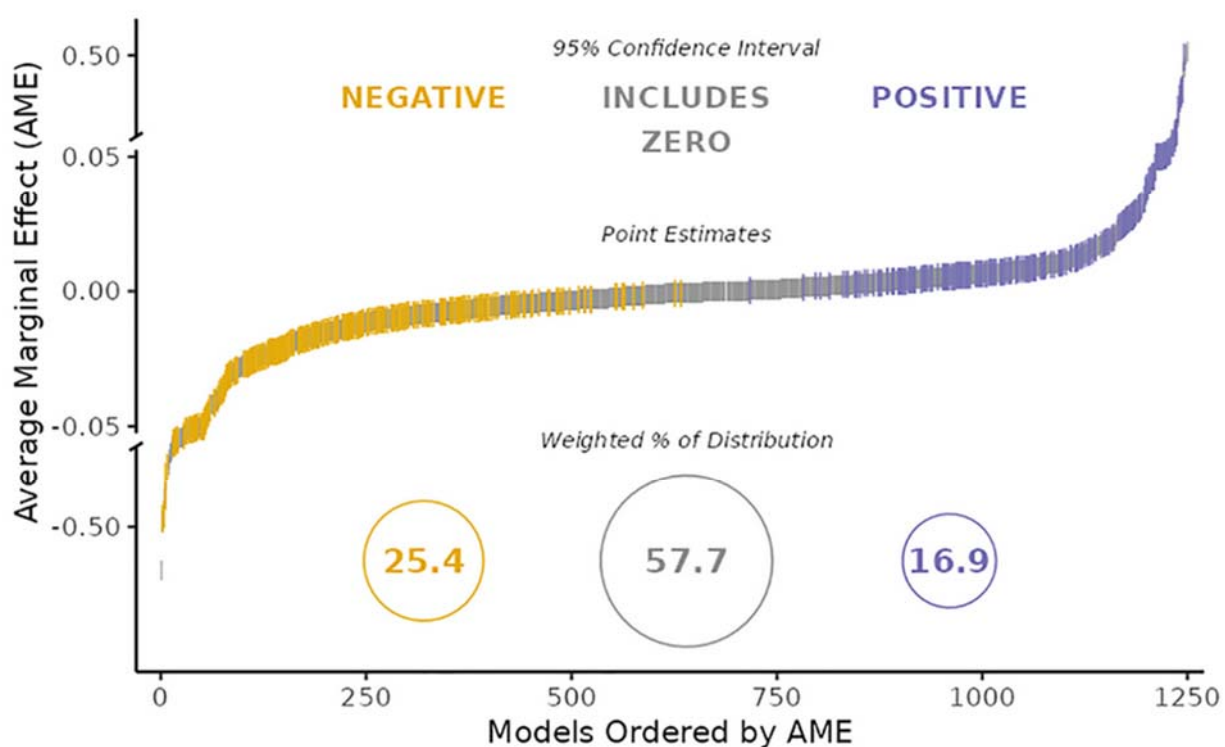
Nate Breznau , Eike Mark Rinke , Alexander Wuttke ,  +162, and Tomasz Żółtak  [Authors Info & Affiliations](#)

Edited by Douglas Massey, Princeton University, Princeton, NJ; received March 6, 2022; accepted August 22, 2022

October 28, 2022 | 119 (44) e2203150119 | <https://doi.org/10.1073/pnas.2203150119>

We coordinated 161 researchers in 73 research teams and observed their research decisions as they used the same data to independently test the same prominent social science hypothesis.

Fig. 1.



Blinding
Double-blinding
Triple-blinding

In practice, however...

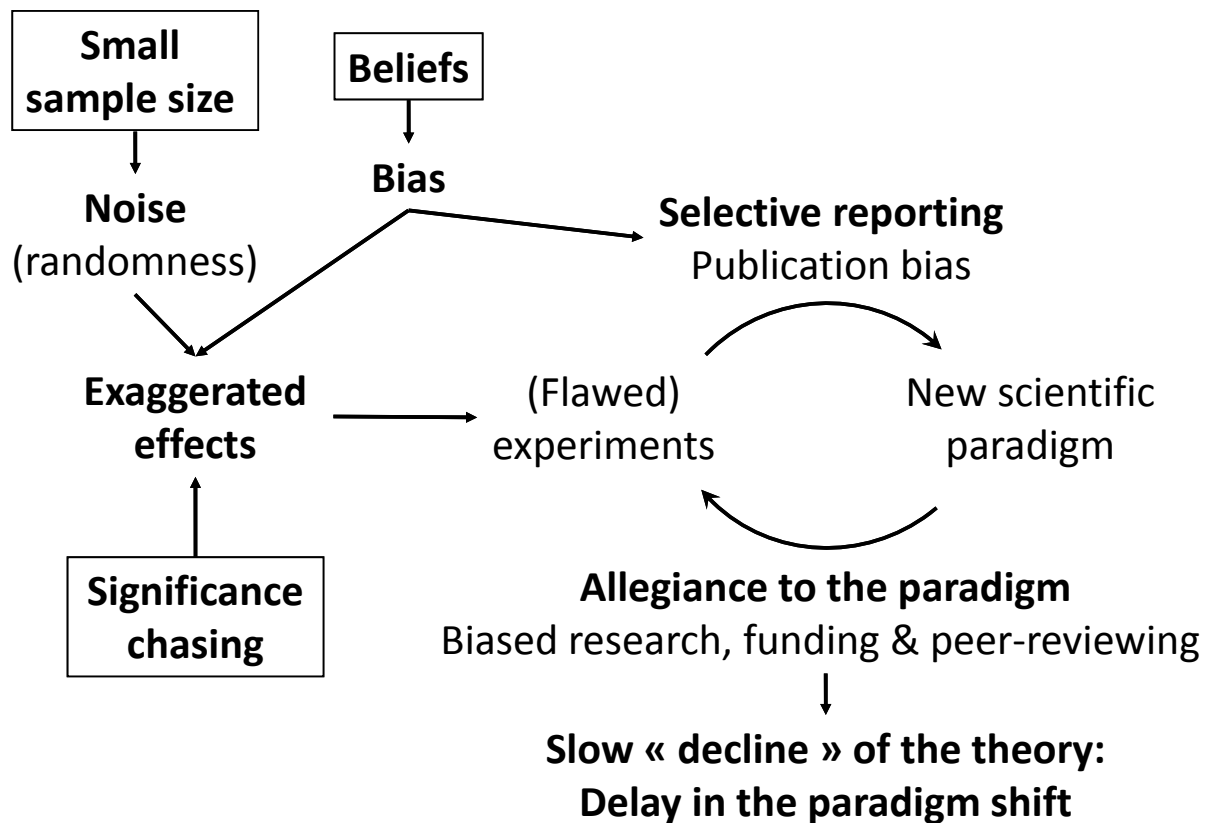
***Never underestimate
confirmation bias***

In practice, however...

***We all suffer from
confirmation bias***

**There are an infinite number of
ways to bias an experiment**

Bias in research & non-replicability : « *decline effect* »



The role of a scientist (and science) is **not** to prove an hypothesis.

It is to try to **demolish it** as much as possible.

If the theory still holds afterwards, then we can label it as **"currently plausible"** (until more data comes in...)

Scientific process, in theory...

*“A scientist is someone who loves
to be proven wrong.”*

“Don't fall in love with your hypothesis.”

– Adriano Aguzzi

Example
Can you find the rule ?

I have chosen a rule that some sequences of three numbers obey and some do not.

Your job is to find out what the rule is using example sequences.

I start by giving you a sequence which obeys the rule:

2

4

8

Obeys the rule

Now it's your turn. Write a number sequence, and I'll tell you whether it satisfies the rule or not. You can test (more or less) as many sequences as you want.

?

?

?

?

When you think you know the rule, describe it in words:

A typical session:

2	4	8	Obeys the rule
1	2	4	Obeys the rule
3	6	12	Obeys the rule
...			

Each number is twice the previous one.

The original rule for the three numbers was that each of them must be larger than the one before.

This game was proposed by the New-York Times (<http://www.nytimes.com/interactive/2015/07/03/upshot/a-quick-puzzle-to-test-your-problem-solving.html>) and is an example of "confirmation bias".

On the NYT website, almost 80% of respondents answered with a rule without trying a "negative control" – a sequence that should not pass their proposed rule. This is an example of confirmation bias, where a person just tries to confirm his/her hypothesis without considering alternative explanations – something that should be avoided, especially in science.

Be careful about the "Bias blind spot"

"cognitive bias of recognizing the impact of biases on the judgment of others, while failing to see the impact of biases on one's own judgment"

https://en.wikipedia.org/wiki/Bias_blind_spot

Experimental design is a Ulysse's pact

https://en.wikipedia.org/wiki/Ulysses_pact



Ulysses and the Sirens, painting by John William Waterhouse

- <https://twitter.com/doctorow/status/1361365103795335169>
- Self-control isn't merely a matter of eliminating your own weaknesses. Self control is primarily about compensating for those weaknesses.
- This manoeuvre has a name: a Ulysses Pact, named for the passage in the Odyssey in which Ulysses pilots his ship through the sirens' sea, eschewing wax-stoppered ears so that he could hear their song, protecting himself by lashing himself to the mast.
- A Ulysses Pact is an act of humility, an admission of frailty

Thank you !

Frederic.Schutz@unil.ch

**Some additional slides,
based on the discussion**

Olfactory exposure to males, including men, causes stress and related analgesia in rodents

Robert E Sorge^{1,2,8}, Loren J Martin^{1,8}, Kelsey A Isbester¹, Susana G Sotocinal¹, Sarah Rosen¹, Alexander H Tuttle¹, Jeffrey S Wieskopf¹, Erinn L Acland¹, Anastassia Dokova¹, Basil Kadoura¹, Philip Leger¹, Josiane C S Mapplebeck¹, Martina McPhail³, Ada Delaney⁴, Gustaf Wigerblad⁴, Alan P Schumann², Tammie Quinn², Johannes Frasnelli^{5,6}, Camilla I Svensson⁴, Wendy F Sternberg³ & Jeffrey S Mogil^{1,7}

We found that exposure of mice and rats to male but not female experimenters produces pain inhibition.

Nature Methods, 11 (6), June 2014. p. 629-632.

nature

International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video |

News & Comment > News > 2017 > November > Article

NATURE | NEWS



Sex matters in experiments on party drug — in mice

Ketamine lifts rodents' mood only if administered by male researchers.

Sara Reardon

17 November 2017

DBP, a Liver-Enriched Transcriptional Activator, Is Expressed Late in Ontogeny and Its Tissue Specificity Is Determined Posttranscriptionally

**Christopher R. Mueller,* Pascal Maire,†
and Ueli Schibler**
Department of Molecular Biology
University of Geneva
CH 1211 Geneva 4
Switzerland

The paper describes a novel D-box-binding transcription factor, which is dubbed it DBP (for albumin site D-Binding Protein).

Cell, 61, 20 April 1990. p. 279-291.

What happens when someone else (in the same lab) tries to reproduce these results ?

What happens when someone else (in the same lab) tries to reproduce these results ?

- After the post-doc (Chris) left, a new PhD student (Jérôme) started and tried to reproduce the results, but couldn't
- The lab head (Ueli) considered retracting the original study, but decided to try the experiment once more.
- ... and he managed to reproduce the results!

What happens when someone else (in the same lab) tries to reproduce these results ?

- They found out that Jérôme (the son of a farmer) was used to wake up early and start his experiments in the morning
- Chris was a night owl, and wouldn't conduct his experiments until the afternoon
- Ueli, as the lab head, was busy during the day, and could only get back to the bench in late afternoon

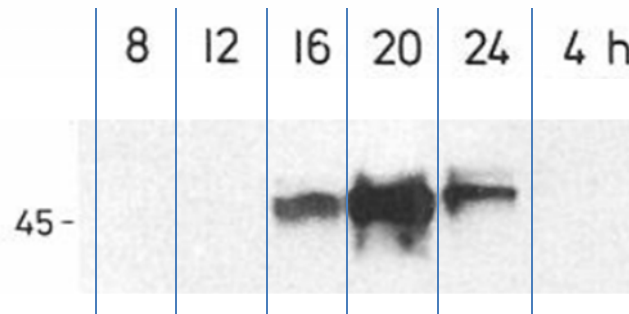
What happens when someone else (in the same lab) tries to reproduce these results ?

- We now know many proteins which follows such a circadian pattern, but DBP is still the one that shows the strongest difference (300x) between the minimum and maximum level.

Expression of the Liver-Enriched Transcriptional Activator Protein DBP Follows a Stringent Circadian Rhythm

Jérôme Wuarin and Ueli Schibler
Department of Molecular Biology
University of Geneva
30 Quai Ernest Ansermet
1211 Geneva-4
Switzerland

Expression according to time of day: 300x higher



Cell, 63 (6), 21 December 1990, p 1257-1266.

“ We got lucky not by carefully planning experiments, but by pure chance, as Jérôme was a morning lark and Chris a night owl. ”

U. Schibler, “Getting Surprising Answers to Unasked Questions”. Cell, 169 (7), 2017, p 1162-1167.

- Veronique Greenwood. *"How the Body's Trillions of Clocks Keep Time"*.
- Quantamagazine, 15 September 2015.
- <https://www.quantamagazine.org/how-the-bodys-trillions-of-clocks-keep-time-20150915>
- also reprinted in Wired, 26 September 2015
- <https://www.wired.com/2015/09/bodys-trillions-clocks-keep-time/>



RESEARCH ARTICLE

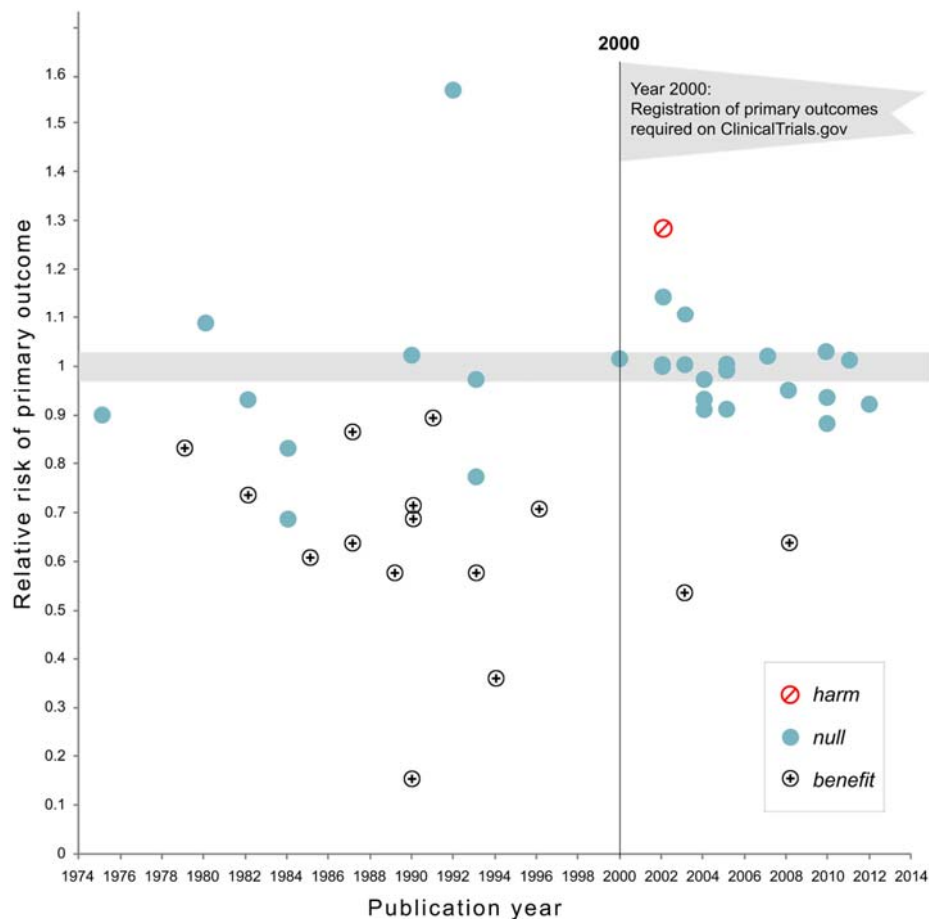
Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time

Robert M. Kaplan^{1*}, Veronica L. Irvin²

1 Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, Rockville, Maryland, United States of America, **2** Oregon State University, Corvallis, Oregon, United States of America

Background

We explore whether the number of null results in large National Heart Lung, and Blood Institute (NHLBI) funded trials has increased over time.



Results

17 of 30 studies (57%) published prior to 2000 showed a significant benefit of intervention on the primary outcome in comparison to only 2 among the 25 (8%) trials published after 2000 ($\chi^2=12.2, df=1, p=0.0005$). There has been no change in the proportion of trials that compared treatment to placebo versus active comparator. Industry co-sponsorship was unrelated to the probability of reporting a significant benefit. Pre-registration in clinical trials.gov was strongly associated with the trend toward null findings.

Conclusions

The number NHLBI trials reporting positive results declined after the year 2000. Prospective declaration of outcomes in RCTs, and the adoption of transparent reporting standards, as required by clinicaltrials.gov, may have contributed to the trend toward null findings.

People and institutions cannot keep their own score accurately. Metrics soon become targets, and are thus gamed, undermined, corrupted, enhanced, misreported, fudged.

Edward Tufte