Reproducibility – a workshop

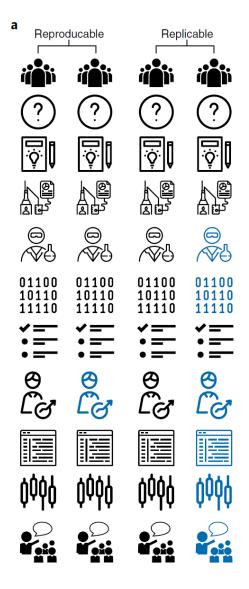
Ulf Toelch

BIF Berlin Institute of Health @Charité

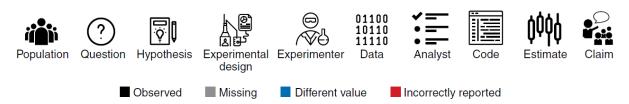


What to expect...

- 1. Reproducibility: Definitions.
- 2. The standard test for truth
- 3. Questionable Research Practices and their effects
- 4. Preregistrations and registered reports.
- 5. Can everything be reproducible?



Reproducible ~ Methods Reproducibility Replicable ~ Results Reproducibility



Patil, P., Peng, R.D. & Leek, J.T. A visual tool for defining reproducibility and replicability. *Nat Hum Behav* **3**, 650–652 (2019). https://doi.org/10.1038/s41562-019-0629-z



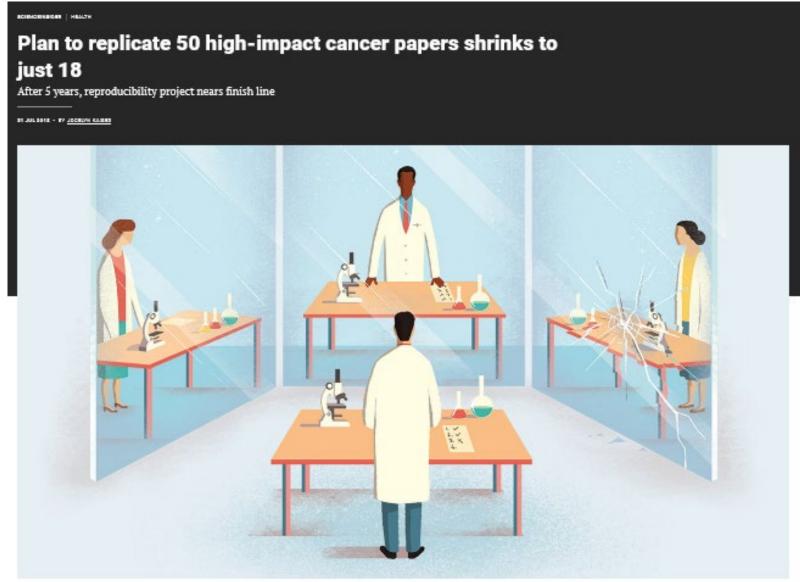
Reproducibility Initiatives

SUBMIT MY RESEARCH LOG IN/REGISTER ABOUT COMMUNITY eLife = HOME MAGAZINE INNOVATION Q Edited by Roger J Davis et al. **Reproducibility Project: Cancer Biology** Investigating reproducibility in preclinical cancer research. f 6 y \times Collection · Dec 10, 2014

 Main findings from 50 high impact citations/publications in cancer research



Results of Reproducibility Project : Cancer Biology





Reproducibility Initiatives

SUBMIT MY RESEARCH LOG IN/REGISTER ABOUT COMMUNITY eLife = HOME MAGAZINE INNOVATION Q Edited by Roger | Davis et al. **Reproducibility Project: Cancer Biology** Investigating reproducibility in preclinical cancer research. f 6 y \times Collection • Dec 10, 2014

 Main findings from 50 high impact citations/publications in cancer research

5 replicated most results

6 replicated parts but not all results

6 were not able to reproduce results

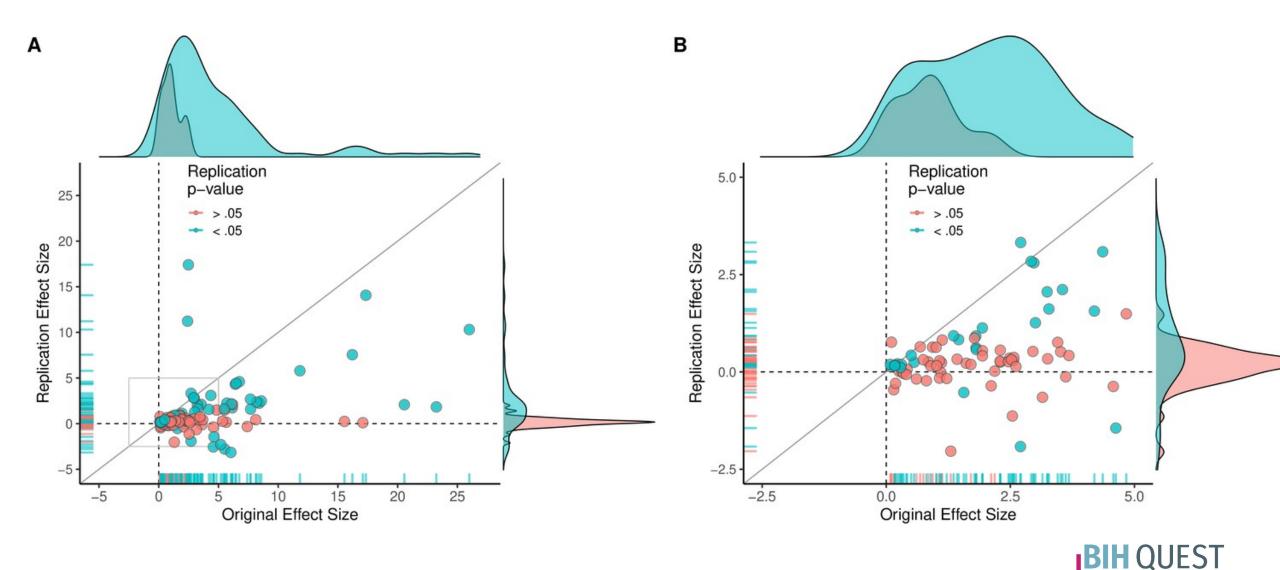


COMPLETED 50 experiments	BARRIERS	
	Modifications implemented Complete Most Most Moderate Some Little No% 25% 50% 75%	IO - N/A
INITIATED 87 experiments	Modifications needed None Few Some Moderate Strong Extrem 0% 25% 50% 75%	ne N/A
	Reagents shared Yes No N/A	100%
	Authors helped	
	Extreme Very Moderate Some Minimal 0% 25%	No 50% 75% 100%
	Protocol clarifications needed Few Some Moderate Strong Extreme	
	0% 25%	50% 75% 100%
DESIGNED	Reagents offered Yes No N/A	
193 experiments	0% 25%	50% 75% 100%
	Code shared Open Yes Some info No N/A	
	0% 25%	50% 75% 100%
	Analysis reported Statistical inference: Test known Test unknown No,	but variation No, but image
	0% 25%	50% 75% 100%
	Data shared Open Raw Summary No	
	0% 25%	50% 75% 100%

Roadblocks to Replication



Replication success limited



Center for Responsible Research

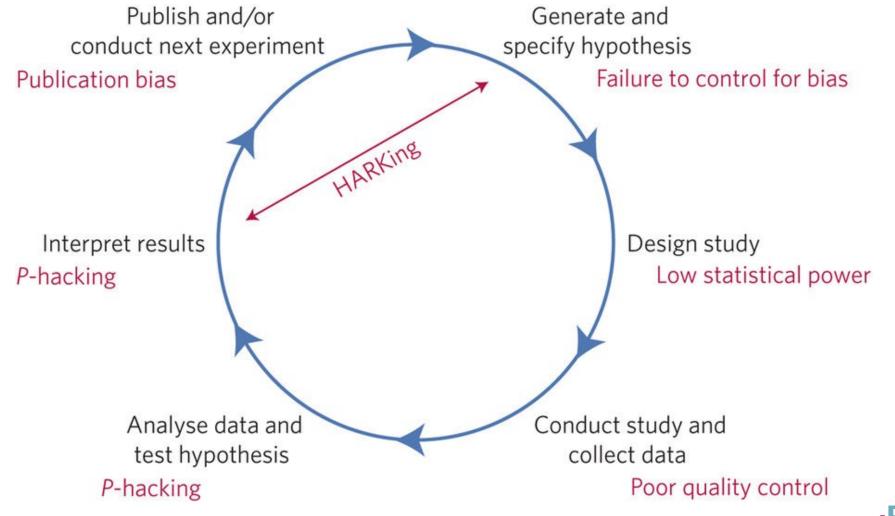
8 12.04.2025 Titel der Präsentation

Science: How to search for the truth...

		WE FOUND NO LINK BETWEEN JELLY BEANS AONE (P > 0.0	1 AND)IH	ITLES THAT. EAR IT'S ONLY ERTAIN COLOR IT CAUSES IT. STS! ANINECOMPT!	
WE POUND NO LINK GETVEEN PRREE ELLY BEANS AND ACKE (P > 0.05).	WE ROUND NO LINK BEIVEN BROWN JELLY BEANS AND ADNE (P > 0.05).	WE ROWID NO LINK BETWEEN PINK JELLY BEING AND ACHE (P > 0.05).	LINK D BLUE	UND NO SETWEEN JELLY AND ACNE 0.05).	WE POWNO NO LINK BETWEEN TEAL JELLY BEARD AND ACHE (P > 0.05).	
WE POUND NO LINK GETWEEN SALINON JELLY BRANS AND ACHE (P > 0.05).	WE FOUND NO LINK GETVEEN RED JELLY BEANS AND ADAE (P > 0.05).	VE FOUND NO LINK BETWEEN TURGUIGE TELLY BEANS AND ACKE (P > 0.05).	LINK (MAGEN BEANS	UND NO SETWEEN ITA JELLY AND ACNE 0.05).	WE FOUND NO LINK BETWEEN YELLON JELLY BEANS AND RAVE (P > 0.05).	
WE ROUND NO LINK BETWEEN GREY JELL BEANS AND ADNE (P > 0.05).	WE FOUND NO LINK BETWEEN TAN JELL BEANS AND ROME (P > 0.05).	VE FOUND NO LINK BOTUEN CYAN JELLY BEARS AND ACKE (P > 0.05).	GREET	UND A SETWEEN N JELLY AND ACHE 0.05).	WE POUND NO LINK BETWEEN MANKE JELLY BEANS AND RAVE (P > 0.05).	
VE ROUND NO LINK BETHEEN BEIGE JELLY BEANS AND ANE (P > 0.05).	WE FOUND NO LINK BETWEEN LINK BETWEEN LINK JELLY BEANS AND ROME (P > 0.05).	WE FOUND NO LINK BETWEEN BAAK JELLY BEANS AND ACHE (P > 0.05).	LINK (PEACH BEANS	UND NO SETWEEN JELLY IND ACKE 0.05).	WE FOUND NO LINK BETWEEN ORANGE JELLY GEANS AND ADNE (P > 0.05).	
GREEN JELLY BEANS LINKED TO ACNE! 95% CONFIDENCE ONLY 5% COMPLEXE SCIENTISE						

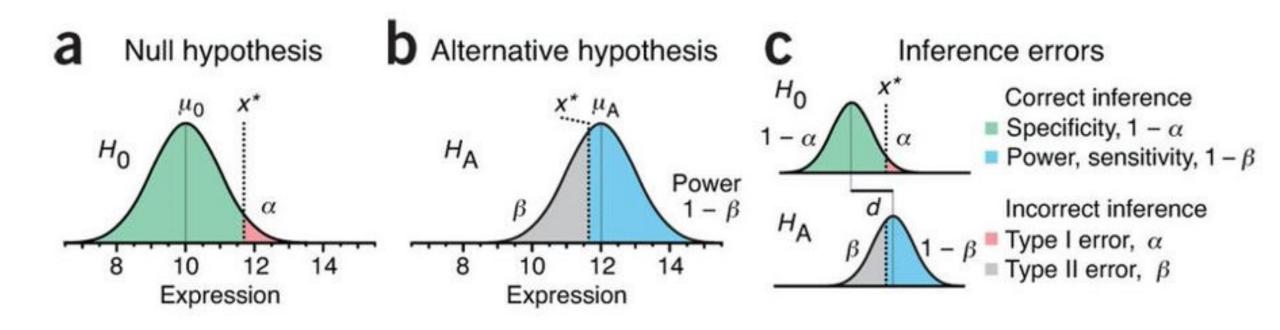
https://xkcd.co m/882/ CC-by BY-NC 2.5 BIH QUEST Center for Responsible Research

Threats to reproducibility





Standard test for the "truth"



https://rpsychologist.com/pvalue/

Krzywinski, M., & Altman, N. (2013). Power and sample size. Nature Methods, 10, 1139.

Standard test for the "truth"

1.Are we doing everything according to plan?2.Can sound published research can be false?



Are we doing everything according to plan? Questionable Research Practices

Questionable research practices (QRPs) are ways of producing, maintaining, sharing, analyzing, or interpreting data that are likely to produce misleading conclusions, typically in the interest of the researcher. QRPs are not normally considered to include research practices that are prohibited or proscribed in the researcher's field (e.g., fraud, research misconduct). Neither do they include random researcher error (e.g., data loss).

Nagy, T., Hergert, J., Elsherif, M., Wallrich, L., Schmidt, K., Waltzer, T., ... & Rubínová, E. (2024, May). Bestiary of Questionable Research Practices in Psychology. https://osf.io/fhk98/download



Team up with your neighbour and find your "favourite" QRP

https://nthun.github.io/qrp-bestiary/qrp_table_wide.html

Explore the effects of some QRPs

https://shiny.psy.lmu.de/felix/ShinyPHack/

Nagy, T., Hergert, J., Elsherif, M., Wallrich, L., Schmidt, K., Waltzer, T., ... & Rubínová, E. (2024, May). Bestiary of Questionable Research Practices in Psychology. https://osf.io/fhk98/download



Parts of preregistrations

Time stamped, read-only version of your research plan

1.Research rationale2.Hypotheses3.Design4.Analytic strategy (+)



http://www.apa.org/science/about/psa/2015/08/pre-registration.aspx

Benefits

Improved use of theory and stronger research methods

A decline in false-positive publications

Reduced File drawer effect

Reduced P-hacking

Reduced HARKing





Pre-registration could lead to undervaluing exploratory research

Rely more on researcher prestige to make decisions about accepting articles for pre-registration.





Think of benefits and concerns you personally see/have when thinking about pre-registration

Write down a list for both

Time: 10 min



Benefits

Improved use of theory and stronger research methods

A decline in false-positive publications

Reduced File drawer effect

Reduced P-hacking

Reduced HARKing

Seven Selfish Reasons for Preregistration:



As predicted



2. Experience the excitement.



4. Profit from online resources.



6. Await your results without fear with in-principle acceptance.



3. Prevent the data from taking you hostage.



5. Increase your reputation and self-image.



 Protect yourself against post-hoc critique.



Pre-registration could lead to undervaluing exploratory research

Rely more on researcher prestige to make decisions about accepting articles for pre-registration.



Confirmation vs Exploration



Preregistration: A Plan, Not a Prison

May 23rd, 2017, Alexander DeHaven

Posted in: Preregistration



https://www.cos.io/blog/preregistration-plan-not-prison

practical examples and solutions: <u>Nosek et al. 2018</u>

- pre-registration:
 - > OSF default template
 - > <u>OSF open-ended</u>
 - > <u>PROSPERO</u>
 - > <u>AsPredicted</u>
 - <u>animalstudyregistry</u>



F

Preregistration for animal experiments



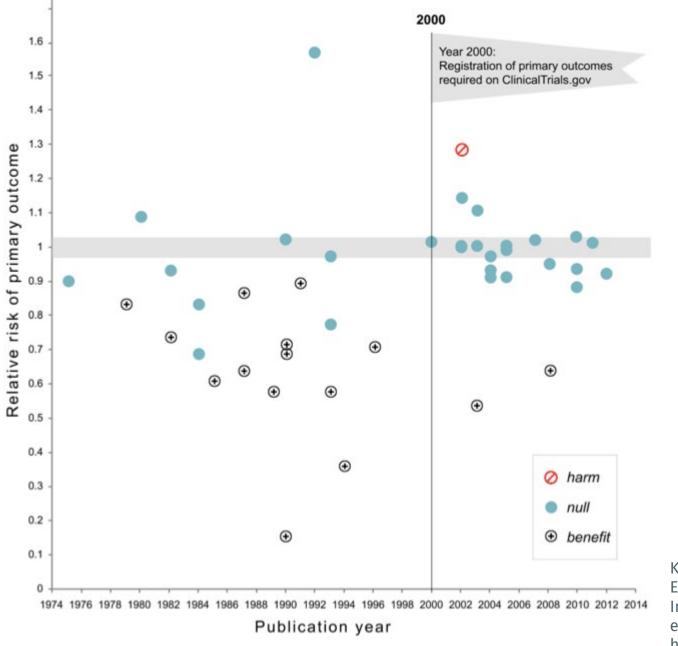
PRECLINICAL TRIALS.EU

Registered Reports



https://cos.io/rr/





Reduced positive findings after mandatory preregistration in clinical trials

Kaplan RM, Irvin VL (2015) Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. PLoS ONE 10(8): e0132382. https://doi.org/10.1371/journal.pone.013238

2



How likely do you think scientific hypotheses are true?



Conducting a power calculation and the Positive Predictive Value



The cure for Alzheimer's

Difficult to understand mechanistically Scientific progress slow

Assumption:

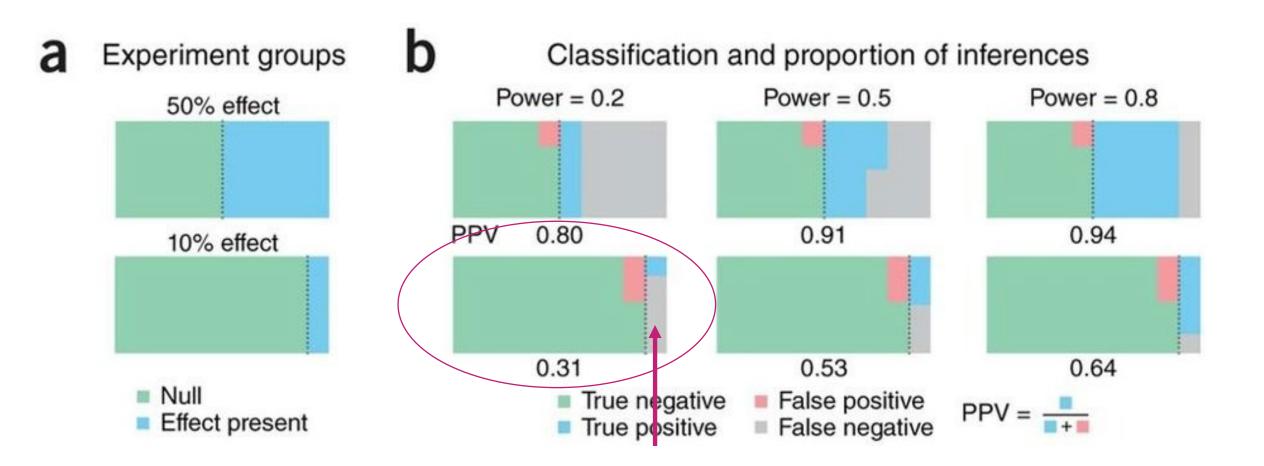
- 1. Low number of our hypotheses are true (<10%)
- 2. Effect sizes will not be overwhelmingly large (Median d=.5)

Typical study: 1 Treatment; 1 Control group N=10 per group

What will be the positive predictive value* after 1000 experiments conducted on different hypotheses?

* Probability that a significant finding reflects a true effect

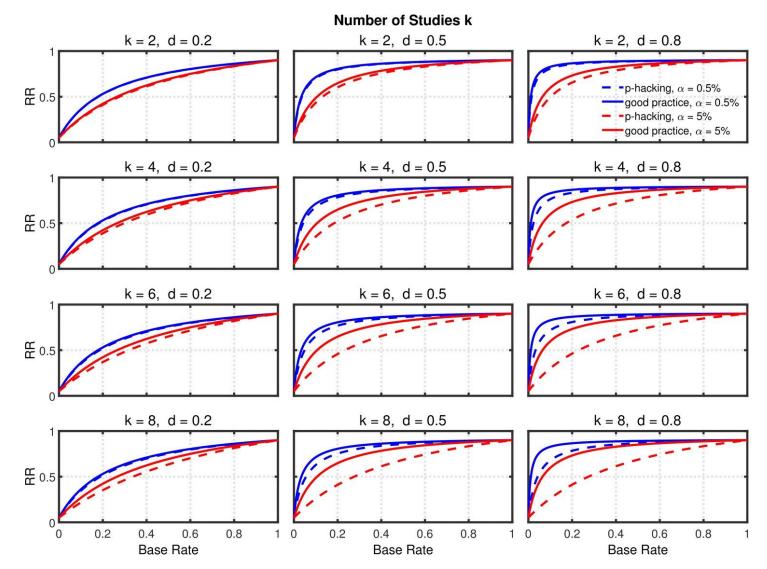




Krzywinski, M., & Altman, N. (2013). Power and sample size. *Nature Methods*, 10, 1139.



The effect of selectively reporting significant studies



Ulrich, R., & Miller, J. (2020). Questionable research practices may have little effect on replicability. *eLife*, *9*, e58237. <u>https://doi.org/10.7554/eLife.58237</u>



We engage in a replication!

We only have 65 studies to replicate. 45 false and 20 real effects

We power these properly at alpha=.05 power=.8 for an effect size of d=.5

PPV=~.9 but we need 64 animals per group!

We have 18 substances to go into clinic out of 1000. But we have overlooked 80 substances that may work!

How are your experiences with replications?



Thank you.



