p-hacking – kein Kavaliersdelikt
p-hacking – not a trivial offense

Coffee Lecture
Nov 19 2019

Eva Furrer
Center for Reproducible Science

What is a p-value?

p-value: chance of seeing the observed results (or more extreme) in the world where the null hypothesis is true

Null hypothesis: There is no difference between two groups, e.g. risk of heart disease is not different for heavy vs moderate coffee drinkers

Null hypothesis significance testing: How ridiculous would it be to believe the null hypothesis is true, given the results?

BUT: Absence of evidence is not evidence of absence

https://www.vox.com/science-and-health/2017/7/31/16021654/p-values-statistical-significance-redefine-0005

Should you calculate a p-value?

In other words: do you have a hypothesis that can be framed into NHST?

Cartoon by Jim Borgman, first published by the Cincinnati Inquirer and King Features Syndicate 1997 Apr 27; Forum section: 1

Exploratory vs. confirmatory research

Exploratory

- small and flexible experiments, different methodologies, including molecular and cellular analyses
- may or may not employ inferential statistics
- goal: catch a small number of promising theories, targets, compounds etc. against a large field

Confirmatory

- rigid and pre-specified designs, a priori stated hypotheses, prolonged durations
- the most clinically relevant assays and endpoints available
- goal: eliminate “false positives” that are captured in exploration.

https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001863

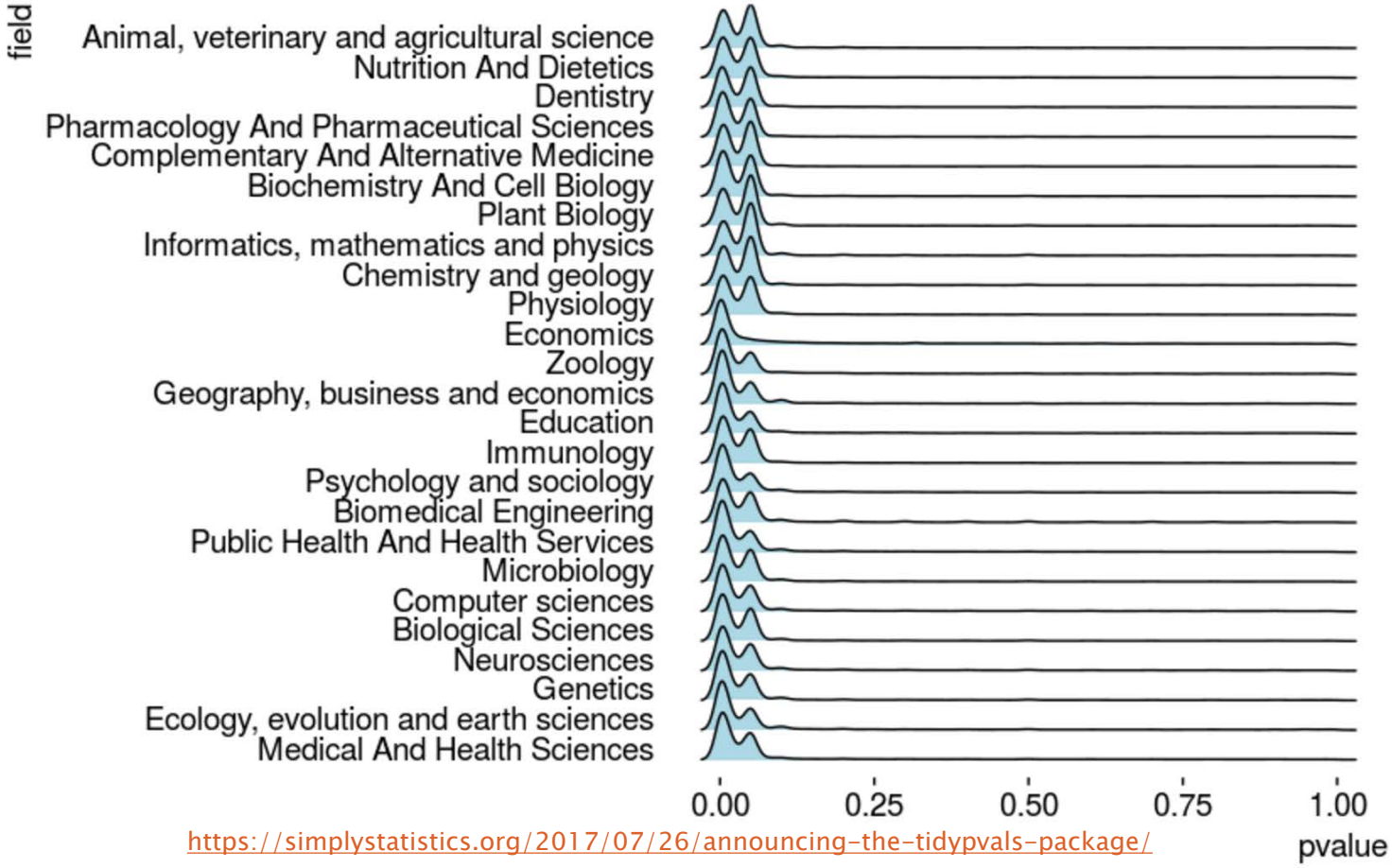
Forms of p-hacking

Inflation of effect sizes through selectively reporting results that are based on result-oriented and unplanned

- decisions to continue data collection
- choice of the „best“ of many recorded response variables
- exclusion or inclusion of outliers
- definition of treatment groups
- exclusion or inclusion of covariates
- stopping of data analysis at significance

Does it happen?

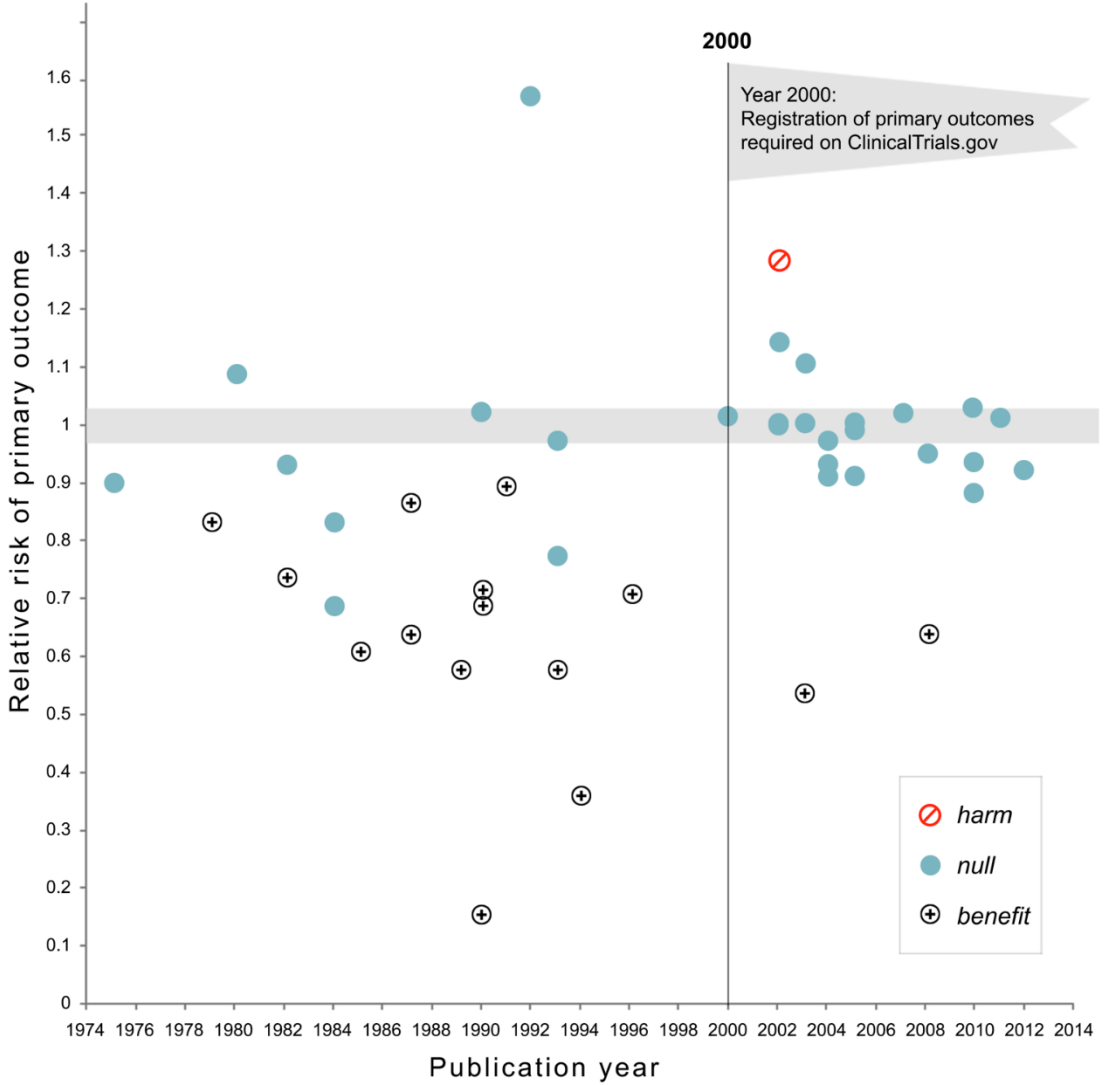
More than 2.5 million p-values across 25 different fields:



https://simplystatistics.org/2017/07/26/announcing-the-tidypvals-package/

Signs of self-correction

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132382



Center for Reproducible Science

Mission

The reproducibility of scientific findings is crucial for the credibility of empirical research. The objective of the Center of Reproducible Science (CRS) is to train the next generation of researchers.

The CRS brings together methodologists from across → University of Zurich (UZH), working in fields which typically do not communicate with each other intensively. This methodological think tank allows to overcome traditional barriers between fields, and aims to determine sound state-of-the-art solutions to scientific challenges.

As a result, UZH researchers who are invested in replication or reproducibility efforts can get together with the methodologists of the CRS either through training activities, direct collaboration, or simply via publications.

See the → current CRS activities here and find → our recent talks here.

You can find → current CRS courses here.

www.crs.uzh.ch


Take home messages

- Absence of evidence is not evidence of absence
- Distinguish exploration from confirmation
- Write and register statistical analysis plans
- www.crs.uzh.ch