

Use of Open Data R Packages to Detect P-hacking in Scientific Literature

Cynthia M. Kroeger, Ph.D.

April 27, 2018



INDIANA UNIVERSITY
SCHOOL OF PUBLIC HEALTH
Bloomington

Critique in Academic Literature

NATURE | COLUMN: WORLD VIEW

Faculty promotion must assess reproducibility

Research institutions should explicitly seek job candidates who can be frankly self-critical of their work, says **Jeffrey Flier**.

12 September 2017

BY MONYA BAKER

Vol. 357, Issue 6353, pp. 759-761
DOI: 10.1126/science.aan4906

RESEARCHERS SURVEYED

| Response | Percentage |
|---------------------------|------------|
| Yes, a significant crisis | 52% |
| Yes, a slight crisis | 38% |
| No, there is no crisis | 3% |
| Don't know | 7% |



Critique in Mainstream Media

The Economist Subscribe Log in or sign up ≡

Problems with scientific research

How science goes wrong

Scientific research has changed the world. Now it needs to change itself



Speaking
Science
Aaron
THE NI
M
can
a pi

By Joel J.

Print edition | Leaders >
Oct 21st 2013

W LOG IN

153

Improvement Efforts

- ◊ Minimizing Spin
- ◊ Correcting Errors within Published Literature
- ◊ Public Deposit of Data and Code
- ◊ Focused Training in Statistics
- ◊ Preregistration of Confirmatory Research
- ◊ Automation of Scientific Processes
- ◊ Technology Development

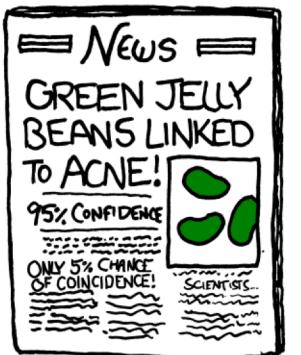
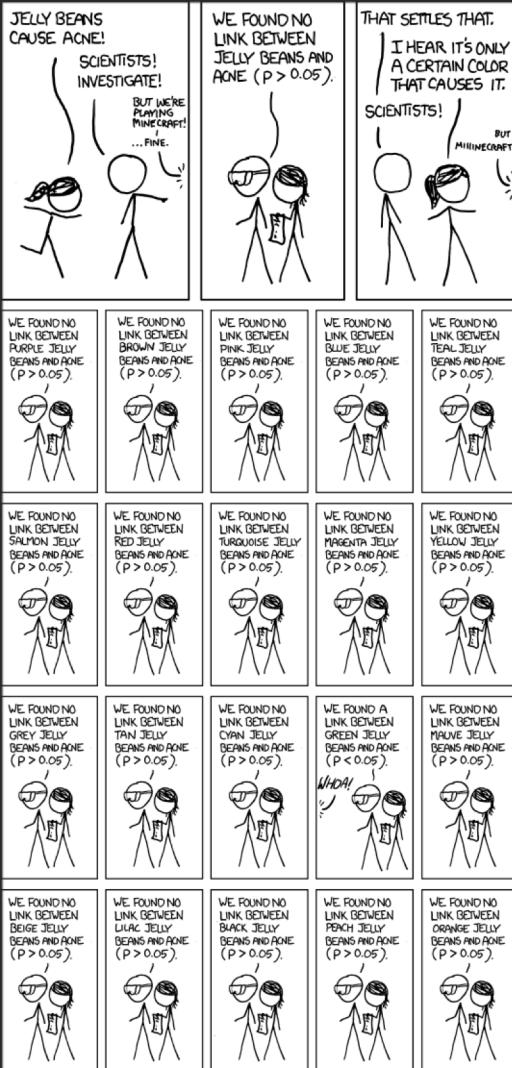
Semi-Automation for Meta-Research

Methods

- ◊ Open Science R Packages (Rentrez)
 - ◆ Replicable Searches
 - ◆ Automatic Data Extraction
- ◊ GitHub
 - ◆ Version Control
 - ◆ Repository for Code/Data
- ◊ Open Science Framework
 - ◆ Study Registration
 - ◆ DOIs for Code & Data

Original Data

- ◊ Detection of P-hacking within Scientific Literature



Detection of P-hacking in Science

P-hacking

- ◊ Data-driven decision making
- ◊ Lack of transparency
- ◊ Bias toward desired p-value
- ◊ Explicit or Implicit

Detection of P-hacking in Science

Problematic?

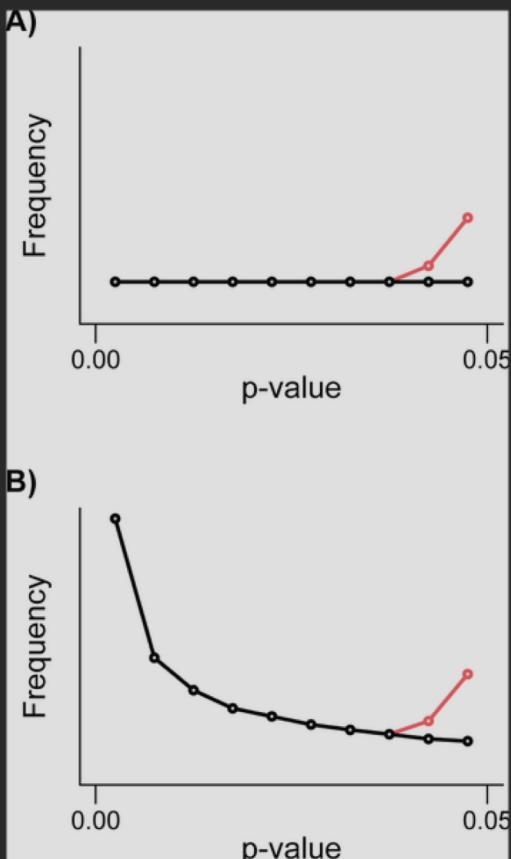
- ◊ Type I error inflation
- ◊ Impedes replicability
- ◊ Undermines credibility
- ◊ Wasteful of resources
- ◊ Indirect harm

Solutions?

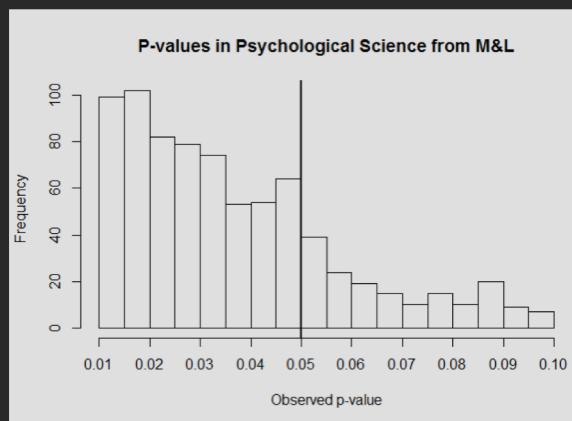
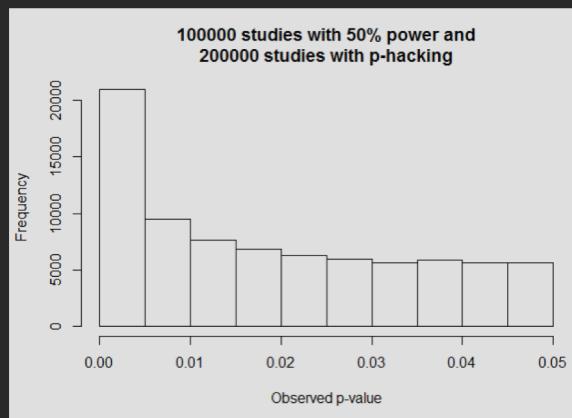
- ◊ Pre-registration
- ◊ Transparency
- ◊ Prevalence/Impact
 - ◆ Survey behavior
 - ◆ Compare protocols to reports
 - ◆ P-value distributions

Detection of P-hacking in Science

The effect of p-hacking on the distribution of p-values in the range of significance

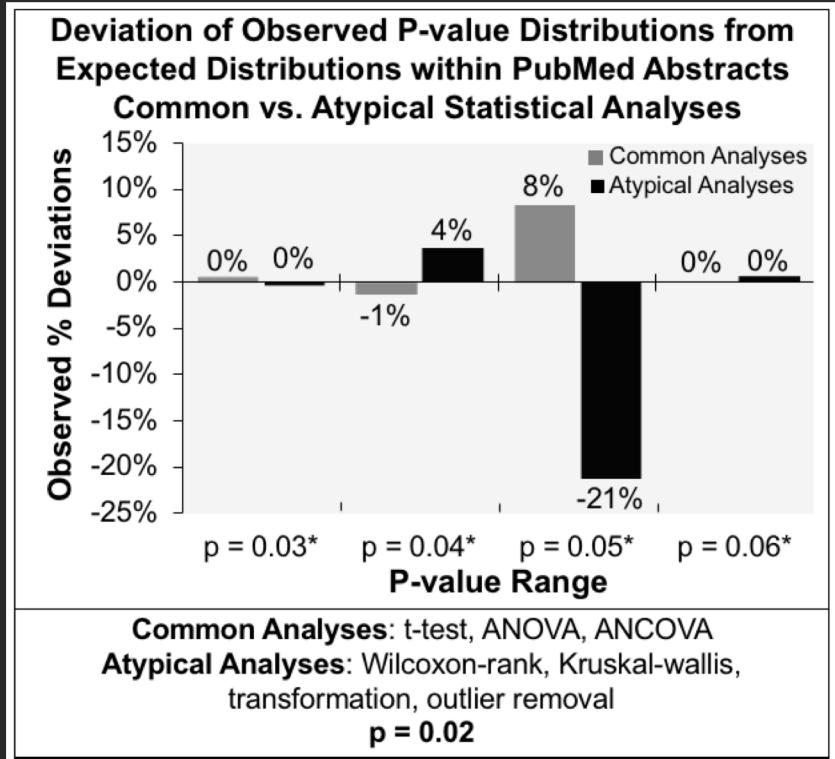


Conditions under which p-hacking does not cause bump and when no p-hacking does



Detection of P-hacking in Science

Specific type of p-hacking: Obtaining a nearly statistically significant result with a “common” test and then switching to an “atypical test” without reporting such



```

# Create objects for individual p-value bin search strings
three <- "(p=.03*[TIAB] OR p=0.03*[TIAB])"
four <- "(p=.04*[TIAB] OR p=0.04*[TIAB])"
five <- "((p 050[TIAB] OR p 0500[TIAB] OR p 05006[TIAB] OR p 0501[TIAB]...
six <- "(p=.06*[TIAB] OR p=0.06*[TIAB])"

# Create objects for exclusive p-value bin search strings
three_ex <- paste(three, "NOT", four, "NOT", five, "NOT", six)
four_ex <- paste(four, "NOT", three, "NOT", five, "NOT", six)
five_ex <- paste(five, "NOT", three, "NOT", four, "NOT", six)
six_ex <- paste(six, "NOT", three, "NOT", four, "NOT", five)
# Create object for common tests search string for Method d
com_d <- "(t-test[TIAB] OR t test[TIAB] OR t-student[TIAB] OR t student[TIAB] OR
anova[TIAB] OR ancova[TIAB])"

# Create object for atypical tests search string for Method d
atyp_d <- "(nonparametric[TIAB] OR non-parametric[TIAB] OR non parametric[TIAB]
OR kruskal-wallis[TIAB] OR rank-sum test[TIAB] OR rank sum test[TIAB]
OR spearman rank correlation coefficient[TIAB] OR
spearman correlation[TIAB] OR wilcox test[TIAB] OR
kolmogorov-smirnov test[TIAB] OR kolmogorov smirnov test[TIAB] OR
u-test[TIAB] OR mann whitney[TIAB] OR mann-whitney[TIAB] OR
wilcoxon-mann-whitney[TIAB] OR wilcoxon[TIAB] OR wilcoxon-rank[TIAB]
OR kruskal wallis[TIAB] OR sign test[TIAB] OR signed-rank[TIAB] OR
friedman test[TIAB] OR mood's median test[TIAB] OR bootstrapping[TIAB]
OR permutation test[TIAB] OR log-transform*[TIAB] OR
log transform*[TIAB] OR log transformed[TIAB] OR
(Outlier*[TIAB] AND remov*[TIAB]))"

# Create object for exclusive common tests search string
com_d_ex <- paste(com_d, "NOT", atyp_d)

# Create object for exclusive atypical tests search string
atyp_d_ex <- paste(atyp_d, "NOT", com_d)

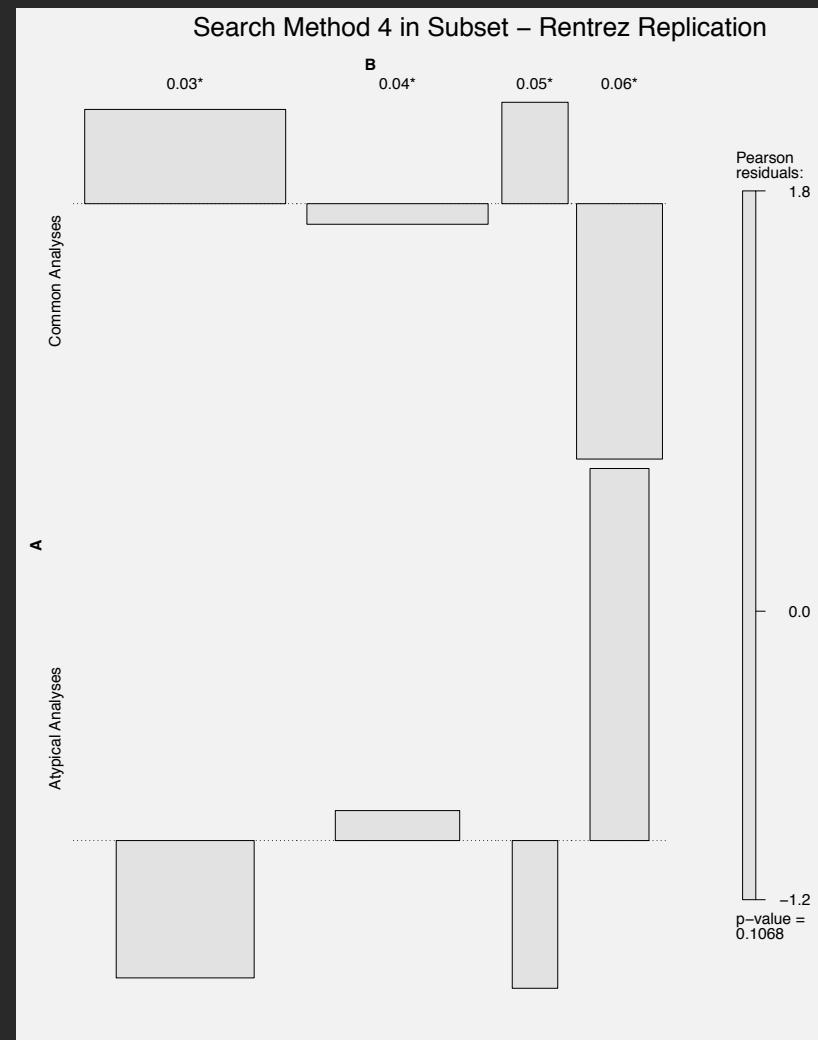
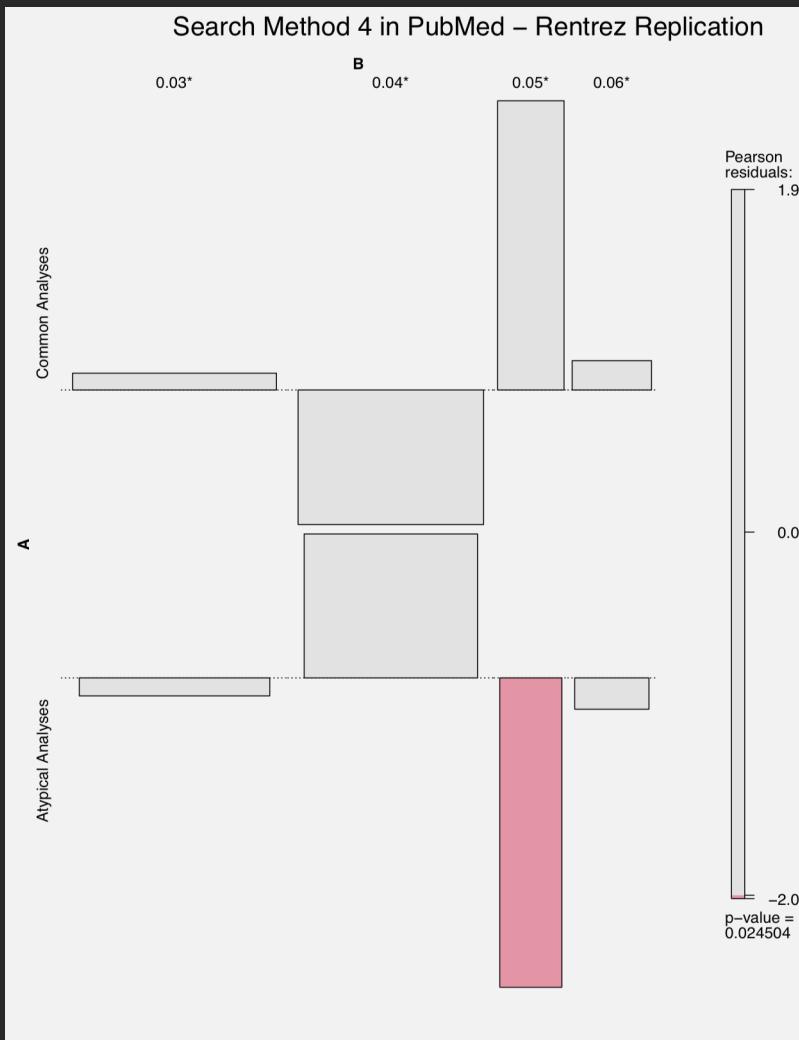
# Create object for each exclusive p-value bin and test type combination search
three_com_d_ex <- paste(com_d_ex, "AND", three_ex)
four_com_d_ex <- paste(com_d_ex, "AND", four_ex)
five_com_d_ex <- paste(com_d_ex, "AND", five_ex)
six_com_d_ex <- paste(com_d_ex, "AND", six_ex)
three_atyp_d_ex <- paste(atyp_d_ex, "AND", three_ex)
four_atyp_d_ex <- paste(atyp_d_ex, "AND", four_ex)
five_atyp_d_ex <- paste(atyp_d_ex, "AND", five_ex)
six_atyp_d_ex <- paste(atyp_d_ex, "AND", six_ex)
# Create counts for each exclusive p-value bin and test type combination
three_com_d_ex_ct <- entrez_search(db = "pubmed",
                                    term = three_com_d_ex,
                                    retmax = 0)

```

Methods

- ◊ Improve search terms
- ◊ Improve process simplicity
- ◊ Improve reproducibility
- ◊ Transparently vet methodology

Results



Discussion

- ◇ Public Health Implications
- ◇ Future directions:
 - Many papers excluded
 - Use fulltext to extract test type chunks
 - Match test type to primary outcome measure

Resources

- ◊ SSRC
- ◊ TurnItIn
- ◊ BITSS
- ◊ Equator Network
- ◊ Center for Open Science
- ◊ Coursera
- ◊ Swirl
- ◊ GitHub Training
- ◊ Software Carpentry
- ◊ OpenIntro Statistics



Acknowledgements

Postdoctoral Mentors

- ◊ David B. Allison, Ph.D.
- ◊ Andrew W. Brown, Ph.D.

Funding

- ◊ NIH/NIDDK: F32DK107157



INDIANA UNIVERSITY
SCHOOL OF PUBLIC HEALTH
Bloomington

Questions?

