

Conference Materials 2018

Cynthia M. Kroeger, Andrew W. Brown, David B. Allison

6/30/2018

Project Description

This project was implemented to further develop methods that were originated previously (*insert DOI for 2014 abstract*) as part of a larger project (F32DK107157) to study common biases and statistical mistakes made in nutrition and obesity research. One such bias is p-hacking. As the talk and poster describe in more detail, this project replicated a previously implemented rapid and high-volume method for detecting p-hacking in scientific literature, as well as refined previous methods. Details of these changes are described in more detail below. This work has since been further developed, and a manuscript is now being prepared for publication.

Conference Description

This work was presented at The American Society for Nutrition (ASN) Annual Scientific Meeting in Boston, MA, 2018.

This work also was presented at the Indiana University School of Public Health Research Day in Bloomington, Indiana, USA, in 2018, as both a talk and a poster.

Contact Information for Investigators

Principal Investigator

- Cynthia M. Kroeger, PhD, Postdoctoral Fellow, Indiana University School of Public Health-Bloomington, Department of Epidemiology and Biostatistics, kroeger@iu.edu

Co-investigators

- Andrew W. Brown, PhD, Assistant Professor Indiana University School of Public Health-Bloomington, Department of Applied Health Sciences, awb1@iu.edu
- David B. Allison, PhD, Dean, Distinguished Professor, Provost Professor, Indiana University School of Public Health-Bloomington, Department of Epidemiology and Biostatistics, allison@iu.edu

Discription of Supporting Materials

Presentation

This work was presented as both a talk and poster.

File name for poster: conference_materials_2018_poster.pdf

File name for talk: conference_materials_2018_talk.pdf

Data

All data were generated within R and thus are produced and described by the code file for this project.

Methodological details

All methodological details are visible and described within the code file for this project.

Reasons for changes in Method 1, compared to previous methods in this repository:

- Common tests: The term “t student” was added for consistency, as the space between words instead of a hyphen yields additional results. The terms “mixed model” and “parametric tests” were removed as the former may be used in atypical contexts and the latter could include abstracts with the phrase “non parametric tests”
- Atypical tests: A librarian was consulted about appropriate search terms for nonparametric tests and the terms “Rank-sum test”, “Rank sum test”, “Spearman Rank Correlation Coefficient”, “Spearman Correlation”, “Wilcox test”, “Kolmogorov-Smirnov Test”, “Kolmogorov Smirnov Test”, “Mann-Whitney U Test”, “Mann Whitney U”, “sign test”, “Signed-Rank”, “friedman Test”, and “Mood’s Median Test” yielded additional relevant results. Further, we found “bootstrapping” and “permutation test” yielded additional relevant results. The term “log transform” was added for consistency, as the space versus hyphen yields additional, relevant results.

Method 2 tests mention of atypical tests in abstracts compared to no mention of these tests in abstracts for each bin and test type.

Analysis code

File name for analysis code: conference_materials_2018_code.R

Study level naming convention: database_method_trial_mode

- database: pm = PubMed; dss = PubMed Dietary Supplement Subset
- method: 1 = method 1; 2 = method 2
- trial: org = original
- mode: ren = rEntrez

Analysis level naming convention: study_function

- study: pm_1_org_ren, dss_1_org_ren, pm_2_org_ren, dss_2_org_ren
- function: xsq = chisq.test; pd = percentage deviation

Code was written by Dr. Kroeger, with assistance from Dr. Brown.

A Note on How to Replicate this Study

Because data were extracted from PubMed via rEntrez on a specific date, one would need to download PubMed history and assess the literature from the point in time noted in the code file. Another option might be to run the search described herein and subtract dates from the date reported in our code file. We have not done this ourselves.

Note on rentrez version: To replicate the figures in these presentations, one must use a rEntrez version less than 1.2. All of the releases are tagged in the GitHub repo. Use devtools GitHub install function with the ‘ref’ argument to set release tag. Otherwise, the search query for the “five” object will be too large a query and will return an error.

Author Contribution

Dr. Kroeger conducted the study, refined the methods, analyzed data, made the presentations, and created all of the open science documents for the repository. Dr. Brown provided editorial assistance with methods and code. Dr. Allison originated the study and is the senior author.

License Information

These materials are licensed under the Creative Commons Attribution Share Alike 4.0.

File name for analysis code: LICENSE.txt

Citation Information

Recommended citation for conference__materials__2018

Please use the following to cite any of the supporting materials herein:

Kroeger CM, Brown AB, Allison DB. Detection of p-hacking in biomedical and dietary supplement scientific literature: Use of open science R packages to improve reproducibility and research transparency. Zenodo. 2018. DOI: *(insert DOI)*.

Recommended citation for published abstract

Please use the following to cite the published abstract:

Kroeger CM, Brown AW, Allison DB. Detection of p-hacking in biomedical and dietary supplement scientific literature: Use of open science R packages to improve reproducibility and research transparency. American Society for Nutrition (ASN): Annual Scientific Meeting. Boston, MA, 2018.

Funding Information

The presentation and development of these materials was supported by the National Institutes of Health (F32DK107157). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or any other organization.