

"Can we run it again, just to be sure?"

"I'm sorry, we can't"

Gabriella Novak

University of Rochester, Program in Statistics

Biostatistics Lab, Johnson Space Center

Mentor: Millennia Young, PhD

"What is reproducible research?"

Results are reproducible when the computational data analysis processes used to generate them can be repeated with consistent input and yield consistent output.

"Is that really necessary?"

Published journal articles retracted because:

- "Inappropriate statistical analysis and inability to reproduce some of the results" [1]
- Data "got terribly mixed" and "published results cannot be reproduced" [2]
- Reanalysis for another purpose revealed that "data reported in this retracted paper were not analyzed properly" [3]
- Concern that the data did not support the figures. Raw data "are not available to allow for a full evaluation of the concerns noted" [4]
- "A calculation error" [5]
- A spreadsheet merging error corrupted the analyzed data set [6]

"But that wouldn't happen to me."

- "Why did I omit those observations?"
- "Where did I get that number from?"
- "How did I make that figure?"
- "Why do the results mentioned earlier in the paper differ from these?"
- "I had incomplete data and now I have to start over again."
- "There was an error so now I have to start over again."

"I'm still unconvinced."

- Increase trust, prevent disaster

"I'm still unconvinced."

- Increase trust, prevent disaster
- Save time and funds

"I'm still unconvinced."

- Increase trust, prevent disaster
- Save time and funds
- Enable continuity and extension of work

"I'm still unconvinced."

- Increase trust, prevent disaster
- Save time and funds
- Enable continuity and extension of work
- Improve quality of scholarship

"I'm still unconvinced."

- Increase trust, prevent disaster
- Save time and funds
- Enable continuity and extension of work
- Improve quality of scholarship
- Better communicate your work

"I'm still unconvinced."

- Increase trust, prevent disaster
- Save time and funds
- Enable continuity and extension of work
- Improve quality of scholarship
- Better communicate your work
- Ease drafting and publishing

"What can I do?"

Stop flirting
with danger

- Data manipulation and analysis in spreadsheet or other graphic-based softwares

"What can I do?"

Stop flirting
with danger

- Data manipulation and analysis in spreadsheet or other graphic-based softwares
- Copy-and-paste



"What can I do?"

Stop flirting
with danger

- Centralize your project files and organize them for clarity and portability

Organize

"What can I do?"

Stop flirting
with danger

- Centralize your project files and organize them for clarity and portability

Organize

- Maintain back-ups and keep them up-to-date

"What can I do?"

Stop flirting
with danger

Organize

Document

- Carry out all data interaction in code recorded in a script



"What can I do?"

Stop flirting
with danger

Organize

Document

- Carry out all data interaction in code recorded in a script
- Document how to execute this code and its intentions



"What can I do?"

Stop flirting
with danger

Organize

Document

- Carry out all data interaction in code recorded in a script
- Document how to execute this code and its intentions
- Record the computational environment the code is meant to operate in



"What can I do?"

Stop flirting
with danger

- Use automatic processes whenever possible

Organize

Document

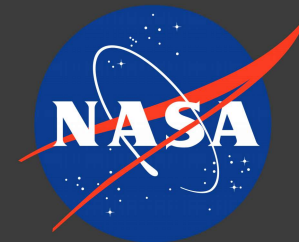
Automate



Acknowledgements

- Millennia Young, PhD, Biostatistics Lab Lead, NASA
- Remainder of the Biostatistics Lab
 - Alan Feiveson, PhD, retiring Biostatistics Lab Lead, NASA
 - Nate Mercaldo, PhD, Biostatistician, KBR
 - Sarah Mercaldo, PhD, LSAH Biostatistician, KBR
- Influential professors and mentors from my university
 - Katherine Grzesik, PhD, Assistant Professor, University of Rochester
 - Joe Ciminelli, PhD, Assistant Professor and Director of the Program in Statistics, University of Rochester





Questions?

References

- [1] Z. Wei, M. Yao, Y. Li, et al. "Retraction Note to: Inhibition of Lipopolysaccharide (LPS)-Induced Inflammatory Responses by Selenium in Bovine Mammary Epithelial Cells in Primary Culture". In: *Inflammation* 38.4 (Aug. 2015), pp. 1662-1662. ISSN: 1573-2576. DOI: [10.1007/s10753-015-0142-0](https://doi.org/10.1007/s10753-015-0142-0). URL: <https://doi.org/10.1007/s10753-015-0142-0>.
- [2] R. Chakrabarti, H. Topf, and M. Schroth. "Retraction Note to: Low Dose Lidocaine for Refractory Seizures in Preterm Neonates". In: *The Indian Journal of Pediatrics* 80.6 (Jun. 2013), pp. 529-529. ISSN: 0973-7693. DOI: [10.1007/s12098-013-0968-0](https://doi.org/10.1007/s12098-013-0968-0). URL: <https://doi.org/10.1007/s12098-013-0968-0>.

References cont.

[3] "NOTICE OF RETRACTION: 'The Emergence of Orthographic Word Representations in the Brain: Evaluating a Neural Shape-Based Framework Using fMRI and the HMAX Model' by Wouter Braet, Jonas Kubilius, Johan Wagemans and Hans P. Op de Beeck. doi:10.1093/Cercor/bhs355, published online November 16, 2012". In: *Cerebral Cortex* 23.8 (Nov. 2012), pp. 2015-2015. ISSN: 1047-3211. DOI: [10.1093/cercor/bhs355](https://doi.org/10.1093/cercor/bhs355). eprint: <https://academic.oup.com/cercor/article-pdf/23/8/2015/1018733/bhs355.pdf>. URL: <https://doi.org/10.1093/cercor/bhs355>.

References cont.

[4] D. Lee, K. H. Lee, H. Park, et al. "Retraction: The Effect of Soluble RAGE on Inhibition of Angiotensin II-Mediated Atherosclerosis in Apolipoprotein E Deficient Mice". In: *PLOS ONE* 9.9 (Sep. 2014). Publisher: Public Library of Science, p. e108684. DOI: [10.1371/journal.pone.0108684](https://doi.org/10.1371/journal.pone.0108684). URL: <https://doi.org/10.1371/journal.pone.0108684>.

[5] M. Lind, L. A. Garcia-Rodriguez, G. L. Booth, et al. "RETRACTED ARTICLE: Mortality trends in patients with and without diabetes in Ontario, Canada and the UK from 1996 to 2009: a population-based study". In: *Diabetologia* 56.12 (Dec. 2013), pp. 2765-2765. ISSN: 1432-0428. DOI: [10.1007/s00125-013-2949-2](https://doi.org/10.1007/s00125-013-2949-2). URL: <https://doi.org/10.1007/s00125-013-2949-2>.



References cont.

[6] S. Kern, I. Skoog, A. Börjesson-Hanson, et al. "Retraction notice to "Lower CSF interleukin-6 predicts future depression in a population-based sample of older women followed for 17 years" [Brain Behav. Immun. 32 (2013) 153-158]". In: *Brain, Behavior, and Immunity* 36 (Feb. 2014), p. 216. ISSN: 0889-1591. DOI: [10.1016/j.bbi.2013.11.011](https://doi.org/10.1016/j.bbi.2013.11.011). URL: <http://www.sciencedirect.com/science/article/pii/S0889159113005424>.