Open Science Panel

Sabine Hoffmann

29.03.2023

Open Science panel

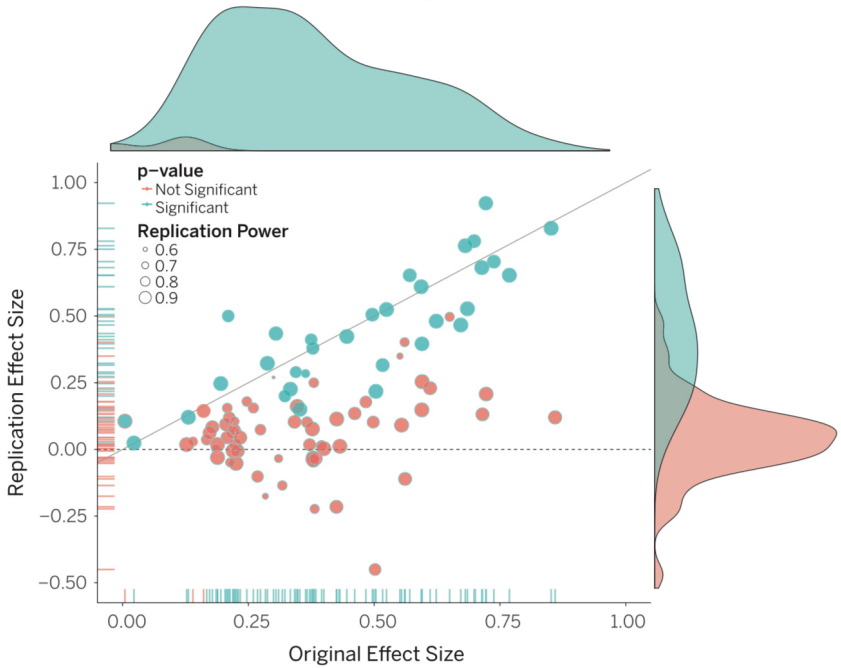
- In the TC on December 15 2022, the Steering group of STRATOS decided to start a new panel on Open Science.
- Chair: Sabine Hoffmann
- Members: Anne-Laure Boulesteix, Daniela Dunkler, Roman Hornung, Michael Kammerer, Kim Luijken, Willi Sauerbrei, Fabian Scheipl, Ewout Steyerberg
- The Open Science panel will write the next Biometric Bulletin article
- New members and a second chair are welcome!

1. What is open science?
2. Which components of open science are important for STRATOS?
3. Current project ideas

What is open science?

The statistical crisis in science

Psychology
[Open Science Collaboration, 2015]



Preclinical research
[Freedman et al., 2015]

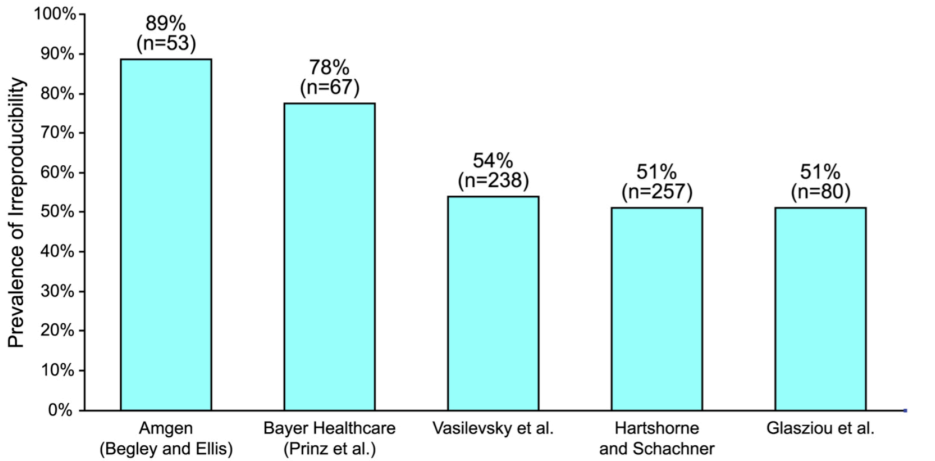


Fig 1. Studies reporting the prevalence of irreproducibility. Source: Begley and Ellis [6], Prinz et al. [7], Vasilevsky [8], Hartshorne and Schachner [5], and Glasziou et al. [9].

The statistical crisis in science

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a *p*-value less than 0.05. Research is not most appropriately represented and summarized by *p*-values, but, unfortunately, there is a widespread

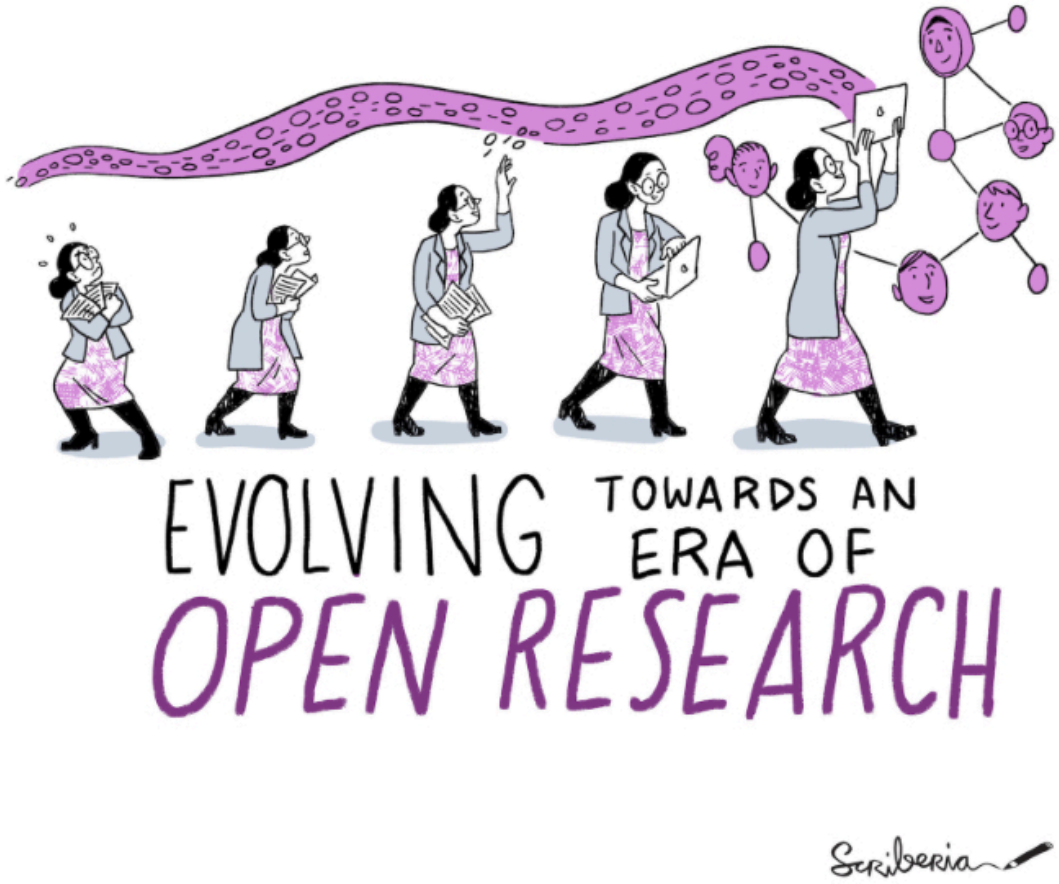
is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability

[Ioannidis, 2005]

[Parsons et al., 2022]: A community-sourced glossary of open scholarship terms

Open science is "an umbrella term that reflects the idea that scientific knowledge of all kinds, where appropriate, should be openly accessible, transparent, rigorous, reproducible, replicable, accumulative and inclusive"

Turing Way definition Open Research



Open research aims to transform research by making it more reproducible, transparent, reusable, collaborative, accountable, and accessible to society.

Which components of open science are important for STRATOS?

Components of open science

- Open access manuscripts

Components of open science

- Open access manuscripts
- Reproducibility:
  - Open code
  - Data availability statement
  - Open data

Components of open science

- Open access manuscripts
- Reproducibility:
  - Open code
  - Data availability statement
  - Open data
- Transparency:
  - Preregistration (e.g. for simulation studies)
  - Complete transparent reporting

Components of open science

- Open access manuscripts
- Reproducibility:
  - Open code
  - Data availability statement
  - Open data
- Transparency:
  - Preregistration (e.g. for simulation studies)
  - Complete transparent reporting
- Replicability:
  - Raising awareness of dangers of result-dependent selective reporting of analysis strategies ("p-hacking", "HARKing" etc.)
  - Dealing with uncertain choices in the analysis of empirical data ("researcher degrees of freedom") while preserving validity of statistical inference

Barriers to open science

- Open access:
  - Open access fees are prohibitive for some journals

Barriers to open science

- Open access:
  - Open access fees are prohibitive for some journals
- Reproducibility:
  - Creating well-structured and well-documented code can be burdensome to authors
  - It can be time consuming to use someone else's code to reproduce paper results

Barriers to open science

- Open access:
  - Open access fees are prohibitive for some journals
- Reproducibility:
  - Creating well-structured and well-documented code can be burdensome to authors
  - It can be time consuming to use someone else's code to reproduce paper results
- Data sharing:
  - Medical datasets often have privacy issues
  - Question of "ownership" of shared data sets and fear of "research parasites"

Barriers to open science

- Open access:
  - Open access fees are prohibitive for some journals
- Reproducibility:
  - Creating well-structured and well-documented code can be burdensome to authors
  - It can be time consuming to use someone else's code to reproduce paper results
- Data sharing:
  - Medical datasets often have privacy issues
  - Question of "ownership" of shared data sets and fear of "research parasites"
- Replicability:
  - Research incentives favour significant and noteworthy results

STRATOS and open science

	STRATOS members	Research community
Open access	Ideally, STRATOS publications should be open access	Underline importance of open access publications
Reproducibility	STRATOS papers should be reproducible	Guidance reproducibility for level 1 audience
	STRATOS papers should use open access data sets	Guidance data sharing while preserving privacy protection and validity of statistical inference
Transparency	Write study protocols (e.g. simulation protocol) and ask for community feedback	Paper illustrating benefits of open simulation setup
Replicability		Guidance dealing with uncertain choices for level 1 audience
		Multi-analyst project?

Current project ideas

Guidance for reproducibility (level 1 audience)

- Guidance paper:
  - Classical approach: Transparent reporting methods
  - Modern approach: Share code and ideally output
  - Outlook: Share data
  - Key points:
    - What is good programming?
    - Compare principles and experiences Biom J [Hofner et al., 2016] and Eur Urol [Assel and Vickers, 2018, Assel et al., 2019a] and ways forward [Assel et al., 2019b, Vickers and Sjoberg, 2022]
- Videos / blog

Paper illustrating benefits of open science practices in methodological research

- Positive example: [Luijken et al., 2022], all code for simulation is available at `https://github.com/Kluijken/CI_CovSel`.
- Other paper illustrating benefits of open science practices in methodological research: [Nießl et al., 2022]

Guidance on how to deal with uncertain choices in the analysis of observational studies (level 1 audience)

- Raise awareness of dangers of result-dependent selective reporting of analysis strategies
- Illustrate uncertain choices in the analysis of research data in biomedicine, potentially through a small and rather informal multi-analyst study
- Propose solutions to deal with these uncertain choices [Hoffmann et al., 2021], tailored to observational studies in biomedical research

Multi-analyst project in biomedical research

- Quantify the impact of all the decisions one is facing during the analysis of observational data to answer one specific research question e.g. if the data are analysed by different analysts
- Many multi-analyst projects in other disciplines, e.g. [Silberzahn and Uhlmann, 2015, Silberzahn et al., 2018, Botvinik-Nezer et al., 2020, Schweinsberg et al., 2021, Hoogeveen et al., 2022]
- Guidance for conducting and reporting multi-analyst studies [Aczel et al., 2021]
- Open methodological challenge: How to make results from different statistical models comparable?

Guidance on data sharing (level 1 audience)

- Important and timely topic for which there is pressing need for guidance on how to share data while preserving
  - privacy protection
  - validity of statistical inference
- Approaches to share privacy protected data:
  - Generation of synthetic data
  - Adding of measurement error and censoring
- For the moment, not much experience with data sharing approaches among members of the panel

Conclusion

- Area in which accessible guidance is urgently needed
- Exciting opportunity to contribute to more reproducible, replicable and credible research

Thank you for your attention!


Aczel, B., Szaszi, B., Nilsonne, G., Van Den Akker, O. R., Albers, C. J., Van Assen, M. A., Bastiaansen, J. A., Benjamin, D., Boehm, U., Botvinik-Nezer, R., et al. (2021).
Consensus-based guidance for conducting and reporting multi-analyst studies.
Elife, 10.


Assel, M., Sjoberg, D., Elders, A., Wang, X., Huo, D., Botchway, A., Delfino, K., Fan, Y., Zhao, Z., Koyama, T., et al. (2019a).
Guidelines for reporting of statistics for clinical research in urology.
BJU international, 123(3):401–410.


Assel, M., Sjoberg, D. D., Catto, J. W., and Vickers, A. J. (2019b).
Innovations in statistical review at european urology.
European urology, 75(1):1.


Assel, M. and Vickers, A. J. (2018).
Statistical code for clinical research papers in a high-impact specialist medical journal.

Annals of internal medicine, 168(11):832.


Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., et al. (2020).
Variability in the analysis of a single neuroimaging dataset by many teams.
Nature, pages 1–7.


Freedman, L. P., Cockburn, I. M., and Simcoe, T. S. (2015).
The economics of reproducibility in preclinical research.
PLoS Biology, 13(6):e1002165.


Hoffmann, S., Schönbrodt, F., Elsas, R., Wilson, R., Strasser, U., and Boulesteix, A.-L. (2021).
The multiplicity of analysis strategies jeopardizes replicability: lessons learned across disciplines.
Royal Society Open Science, 8(4):201925.


Hofner, B., Schmid, M., and Edler, L. (2016).

Reproducible research in statistics: A review and guidelines for the biometrical journal.
Biometrical journal, 58(2):416–427.


Hoogeveen, S., Sarafoglou, A., van Elk, M., and Wagenmakers, E.-J. (2022).
Many-analysts religion project: Reflection and conclusion.


Ioannidis, J. P. A. (2005).
Why most published research findings are false.
PLoS Med, 2(8):e124.


Luijken, K., Groenwold, R. H., van Smeden, M., Strohmaier, S., and Heinze, G. (2022).
A comparison of full model specification and backward elimination of potential confounders when estimating marginal and conditional causal effects on binary outcomes from observational data.
Biometrical Journal.


Nießl, C., Hoffmann, S., Ullmann, T., and Boulesteix, A.-L. (2022).

Explaining the optimistic performance evaluation of newly proposed methods: a cross-design validation experiment.
arXiv preprint arXiv:2209.01885.


Open Science Collaboration (2015).
Estimating the reproducibility of psychological science.
Science, 349(6251):aac4716.


Parsons, S., Azevedo, F., Elsherif, M. M., Guay, S., Shahim, O. N., Govaart, G. H., Norris, E., O,Äômahony, A., Parker, A. J., Todorovic, A., et al. (2022).
A community-sourced glossary of open scholarship terms.
Nature Human Behaviour, 6(3):312–318.


Schweinsberg, M., Feldman, M., Staub, N., van den Akker, O. R., van Aert, R. C., Van Assen, M. A., Liu, Y., Althoff, T., Heer, J., Kale, A., et al. (2021).
Same data, different conclusions: Radical dispersion in empirical results when independent analysts operationalize and test the same hypothesis.

Organizational Behavior and Human Decision Processes, 165:228–249.


Silberzahn, R. and Uhlmann, E. L. (2015).
Crowdsourced research: Many hands make tight work.
Nature, 526(7572):189–91.


Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., et al. (2018).
Many analysts, one data set: Making transparent how variations in analytic choices affect results.
Advances in Methods and Practices in Psychological Science, 1(3):337–356.


Vickers, A. J. and Sjoberg, D. D. (2022).
Methods modernizing statistical reporting in medical journals: Challenges and future directions.
European Urology, pages S0302–2838.


Vickers, A. J., Sjoberg, D. D., et al. (2015).
Guidelines for reporting of statistics in european urology.

Eur Urol, 67(2):181–187.

Wagenmakers, E.-J., Sarafoglou, A., and Aczel, B. (2022).
One statistical analysis must not rule them all.
Nature, 605(7910):423–425.