

Editor Decision Support Tools

MODERATOR:
Tony Alves

Aries Systems Corporation
North Andover, Massachusetts

SPEAKERS:
Elizabeth Caley

Meta, Chan Zuckerberg Initiative
Toronto, Ontario, Canada

Anita Bandrowski

SciCrunch/NIF/RRID
University of California,
San Diego
La Jolla, California

Timothy Houle

Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Chadwick DeVoss

NEX7, StatReviewer
Madison, Wisconsin

REPORTER:
Darren Early

American Society for Nutrition
Rockville, Maryland

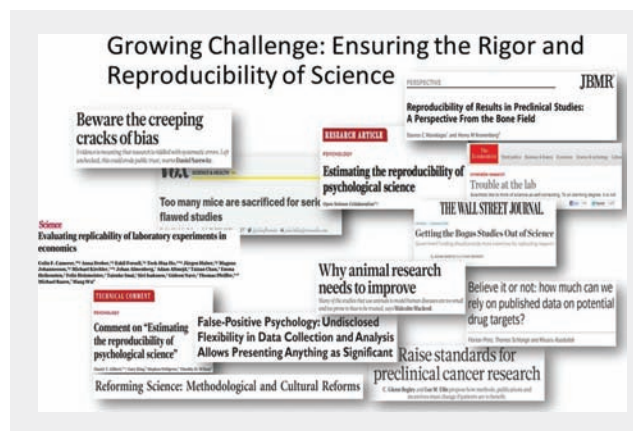


Figure 1. Is reproducibility really a problem?

Tony Alves introduced the session by informing the audience he would focus on three new tools: Meta, the Resource Identification Initiative, and StatReviewer. Elizabeth Caley began by noting that Meta had recently been acquired by the Chan Zuckerberg Initiative, which strives to develop collaborations between scientists and engineers, enable tools and technologies, and build support for science. The Meta Science platform was built using artificial intelligence to enable article discovery. It is currently used by researchers at >1200 institutions and includes 44 million unique pages. The Bibliometric Intelligence tool uses deep predictive profiling to predict Eigenfactor, citations, and top percentile rank in order to answer three core questions about a submitted manuscript: 1) Is it a fit for the journal? 2) What is its potential impact? and 3) Who are the best reviewers for it? This analysis helps editors pinpoint manuscripts at the time of submission that are appropriate for their journals and likely to be of high impact. Bibliometric Intelligence can thus be used to pre-rank manuscripts and intelligently cascade them to sister journals within a publisher's portfolio. The algorithm's results are regularly tested against the actual performance of articles. A detailed white paper on how Bibliometric Intelligence works can be found online.¹

Anita Bandrowski opened her presentation by asking whether reproducibility was really a problem (Figure 1). She noted in a *Nature* survey of >1500 researchers, 90% of respondents indicated there was a reproducibility crisis. Among the four areas the NIH Rigor and Transparency in Research project noted as needing more attention was authentication of key biological/chemical resources. Key biological resources that are the most common failure points in experiments include antibodies, cell lines, and organisms (transgenic). The cost of having key biological resources fail and result in irreproducible results was estimated by Freedman et al. at >\$28 billion annually.² A significant part of the problem (\$10.2 billion annually) is that researchers do not adequately

identify the key biological resources used. This is where Research Resource Identifiers (RRIDs) can be helpful. Other problems include papers published with contaminated cell lines or antibodies that do not work. In both cases, researchers can search the Resource Identification Portal³ for cell lines or antibodies and be notified if problems are known. An RRID author workflow has been developed in which a journal directs an author to the RRID portal, and the author then searches for an antibody and copies "Cite This" text into the manuscript; when the paper is published, the RRID appears so that resources are identifiable. An RRID shortcut called SciScore is also being developed whereby a journal workflow directs a paper's methods section to SciScore, which generates a report for the author, who then uses links to add RRIDs to the paper. RRIDs are found in hundreds of journals, and the adoption of this simple, elegant system is growing rapidly.

The final presentation was given by Timothy Houle and Chadwick DeVoss. In explaining the need for StatReviewer, they noted the poor quality of much published medical research, the low statistical power and errors found in many published articles, and the hypothesis by John Ioannidis that most research findings are actually false. These problems arise because 1) most peer reviewers are not professional methodologists or statisticians, 2) most authors are not also experts in statistics or methodology, 3) only 33% of journals employ a professional statistician, and 4) many scientists, especially the best ones, do not have sufficient time for conducting peer review. StatReviewer provides a solution in automating elements of the statistical and methodological review process, including reporting guidelines such as Consolidated Standards of Reporting Trials (CONSORT) and Strengthening the Reporting of Observational Studies in Epidemiology (STROBE), general statistical reporting, and

CONTINUED

uniform requirements for medical journals. StatReviewer features four applications: parsing, categorizing (determining manuscript type), scanning, and reporting. Analyses provided by StatReviewer include both statistical reporting (what was done and what was found) and statistical design (is the test appropriate?). Coverage of the statistical reporting component is 100%, whereas only 20% of the statistical design component is covered; thus, StatReviewer is meant to supplement, rather than replace, a human reviewer.

References

1. Yang L, Vembu S, Adawi A, Shai O. Enabling editors through machine learning: examining the data science behind Meta Bibliometric Intelligence. Available from: https://medium.com/@meta_6493/enabling-editors-through-machine-learning-81b528b496ce
2. Freedman LP, Venugopalan G, Wisman R. Reproducibility2020: progress and priorities. *F1000Research*. 2017;6:604. <https://doi.org/10.12688/f1000research.11334.1>
3. Research Identification Portal. Available from: <https://scicrunch.org/resources>