J Biosci (2022)47:15
DOI: 10.1007/s12038-022-00259-6

Editorial

Reproducibility and replicability in science: A Sisyphean task

The world received a confidence booster in the power of the scientific method, having witnessed and participated in the recent development of successful vaccines against SARS-COV-2. The world also got a peek into scientific controversies, the clamour for more transparency and data sharing, besides the requirement for rigorous testing, adequate sample sizes, false positives, false negatives, risk probabilities, and population variation. For an interested lay person, or even for a practising scientist, this was the equivalent of a crash course on the world stage on how science is done, warts and all, but where science triumphed in the end.

Behind the scenes, science is in a maelstrom, confronting many demons, and the Better Angels of our Nature are struggling. What is the truth and how does one find it? Are there many truths? Scientists need to solve problems associated with the practice of science, and the sooner the better.

Two major issues facing science today are those of reproducibility and replicability of results. Vigilance on both matters has led to a spate of retractions of papers, e.g. a total of 508 retractions from laboratories in India in the biomedical sciences (Elango 2021). The US National Academy of Sciences, Engineering and Medicine Report (2019) has defined reproducibility in the very narrow context of computer science, requiring all code and data be presented in a paper with adequate transparency and explanation such that anyone running the code would get the same results. This narrow view of reproducibility apparently springs from the work of the seismologist and computer scientist Jon Claerbout (Goodman *et al.* 2016). A broader definition of reproducibility would entail using the same data and methods and producing the same results (Hillary and Rajtmejer 2021). The US Academy justified the narrow sense definition of reproducibility by the current era of big data and burgeoning computer programmes, such that a call for greater data sharing and transparency in data analyses and open-source code are warranted.

Replicability as defined by the US Academy document is the probability of getting the same results using a different data set, but maintaining the same protocols as were published, such as confirming the efficacy of an anti-cancer drug in a different laboratory and the same human cell lines, ordered from the same source, or requested from the authors. There is also the factor of generalisability, which is whether the same drug will deliver the same results on different cell lines. A report published by the Netherlands Academy of Sciences (KNAW 2018) defines reproducibility differently. In this report, reproducibility is the extent to which a replication study's results agree with those of the earlier study. There is therefore confusion in the definition of the terms themselves, although everyone believes that they know what is being said. The Netherlands Academy has gone so far as to set aside €3 million for the first set of projects on replication studies, and has declared that replication must become an essential component of all scientific research.

An attempt to replicate the results in high-ranking journals in the areas of economics, psychology and general science shows that non-replicable studies are cited more often than those that can be replicated (Serra-Garcia and Gneezy 2021). The reasons for this are still not clear, and whether papers that make more novel but non-replicable claims are more likely to be cited needs to be examined. There are instances of influential papers starting a trend of research which was misled by the lack of veracity in the trend-setting publications. This has plagued the field of ecology to the extent that it has given rise very recently to the setting up of SORTEE (Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology). It will be interesting to watch how this Society develops. It is

http://www.ias.ac.in/jbiosci
Published online: 25 February 2022

currently manned by idealistic and energetic students and early career researchers who believe that the truth must out.

Yusuf Hannun, Director of Stony Brook University's Cancer Centre in New York, says that one of the "biggest frustrations as a scientist [is] that it is so hard to know which exciting results are sturdy enough to build on" (Hannun 2021). Hannun blames some of this unfortunate situation on inadequate training which has resulted often in wasted resources, and champions the need for a set of replicated experiments to be incorporated into all schools of science so that young practitioners are made aware of the need for careful replication and the rigour of doing science. Another important reason for such hastily reported and often unverified research is the pressure to continually publish large numbers of papers so that stories are carved up into disjointed elements where the overall question is obscure, and the data themselves are inadequately authenticated. Richard Harris (2017), a respected science journalist, has written a hard-hitting book on wasted financial resources in the pursuit of sloppy and ill-advised science.

One clear hurdle for reproducibility is authenticity of data. Furthermore, what is considered biologically significant is certainly not the same as being statistically significant. There are many pitfalls; replication studies and the explicit requirement for more rigour will solve many of them, especially perhaps curbing some of the reasons for lack of replicability. These include genuine errors; intentional deception; mistakes due to accelerated process without adequate quality control measures; errors that come from large collaborations, where it is impossible to individually vouch for all sets of data; interdisciplinary research where all scientists are not aware of the limitations and quality of different data sets; and finally intentional manipulation or misinterpretation of data due to political interference.

Should we then expect less of a scientific paper (Camerer *et al.* 2018; Amaral and Neves 2021; Camerer *et al.* 2018)? Should we skim a scientific paper just for ideas, and accept that many of the claims may not bear up under scrutiny? Would we have to examine the scientific pedigree of the authors and hope that a good scientific culture has been inculcated, and thereby have more confidence in the findings?

How can results be standardised? Which results should be standardised? Is the inherent variation in the system sufficiently interesting that it too needs to be documented and explored so that causal mechanisms are understood in a more nuanced and context-dependent manner (Bryan *et al.* 2021)? In the rush to standardise, are we losing out on population-level variability and response? I submit that more science academies the world over should grapple with these demons and come up with good measures for grounding research findings in solid and verifiable reality.

In starting out as the Chief Editor of *Journal of Biosciences*, I took heart from May Berenbaum's Editorial when she began recently as Chief Editor of *PNAS*. The editorial is titled: "On zombies, struldbrugs, and other horrors of the scientific literature" (Berenbaum 2021). I leave it to you to find out what her horrors are, and to imagine what mine might be.

Journal of Biosciences is committed to publishing papers of broad scientific interest, and those that adhere strictly to the principles of scientific ethics and good scientific practice. Plagiarism identification software has reduced the embarrassment of pilfered text, but only vigilant editors, reviewers and readers can identify pilfered ideas. A robust network of sensitised professionals is what *Journal of Biosciences* will depend on to bring the best science with the fine attributes of reproducibility and replicability into its ambit.

References

Amaral OB and Neves K 2021 Reproducibility: expect less of the scientific paper. *Nature* **597** 329–331

Berenbaum MR 2021 On zombies, struldbrugs, and other horrors of the scientific literature. *Proc. Nat. Acad. Sci. USA* **118** e2111924118

Bryan CJ, Tipton E and Yeager DS 2021 Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nat. Hum. Behav.* **5** 980–989

Camerer CF, Dreber A, Holzmeister F, *et al*. 2018 Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nat. Hum. Behav.* **2** 637–644

Elango B 2021 Retracted articles in the biomedical literature from Indian authors. *Scientometrics* **126** 3965–3981

Goodman SN, Fanelli D and Ioannidis JP 2016 What does research reproducibility mean? *Sci. Transl. Med.* **8** 341ps12

Hannun YA 2021 Build a registry of results that students can replicate. *Nature* **600** 571–571

Harris R 2017 *Rigor Mortis: How Sloppy Science Creates Worthless Cures, Crushes Hope, and Wastes Billions* (New York: Basic Books)

Hillary F and Rajtmajer S 2021 How failure to falsify contributes to the replication crisis in the clinical neurosciences. *https://www.psyarxiv.com*

KNAW (Royal Netherlands Academy of Arts and Sciences) 2018 *Replication Studies: Improving Reproducibility in the Empirical Sciences*

National Academies of Sciences, Engineering, and Medicine 2019 *Reproducibility and Replicability in Science* (Washington, DC: The National Academies Press). *https://doi.org/10.17226/25303*

Serra-Garcia M and Gneezy U 2021 Nonreplicable publications are cited more than replicable ones. *Sci. Adv.* **7** eabd1705

Renee M Borges

Indian Institute of Science, Bangalore, India

(Email, renee@iisc.ac.in)

Editor-in-Chief, Journal of Biosciences