

Written Evidence Submitted by the Institute for Scientific Information, Clarivate (RRE0077)

Summary of evidence

Upholding research integrity is a joint effort performed by research institutions, researchers, funders, publishers, and database providers. Each must take steps to ensure the validity, accuracy, and overall authenticity of scholarly output. It is also a global endeavor which cannot be solved by one country alone.

While the primary responsibility for research integrity and reproducibility may lie with researchers, they need to be supported by their institutions, publishers, funders and others. Institutions in particular must integrate qualitative indicators for promotions and tenures and move away from using single metrics. This will remove the unnecessary pressure to publish for the sake of publishing. Quality should always trump quantity, especially in science.

A variety of policies and schemes to promote ethical research and reproducibility already exist. In our experience effective policies or schemes enacted to encourage reproducible research consider the value of diversity of thought and experience, should be independent from any one stakeholder, use a balance of quantitative and qualitative methods and understand the importance of partnership across the ecosystem.

Systemic changes needed to deliver a positive impact on academia's approach to reproducible research:

1. Stop misuse of journal-level metrics as a proxy for researcher and article-level evaluation
2. Include publications in all high-quality journals – not just high impact journals – for research assessment⁽¹⁾
3. Reduce use of single point metrics for research assessment⁽²⁾
4. Promote adoption of open research practices and incentives to share research outputs along the entire process, rather than just the final published article
5. Increase collaboration and intelligence sharing between stakeholder groups – reproducibility & research integrity is a shared responsibility
6. Introduce accredited training in research integrity as a pre-requisite to being awarded a PhD or grant funding

The integrity and reproducibility of research is monitored and validated by the research community through the process of peer review. This is applied to project proposals, reports and through editorial review to everything that is submitted for publication before it is published as an accepted part of the research corpus.

Clarivate is owner of the Web of Science™, which is recognised as the longest established and most authoritative index of research literature and is the primary international source of data for policy and academic analysis of research activity and achievement. As a respected research partner to innovators across the world - our data, insights and analysis are used by >95% of the world's top research institutions, multiple governments and national research agencies for research evaluation and assessment - we are consequently experienced in monitoring the significance of global research literature, since this determines the relevance, quality and value of our systems and services. We are deeply concerned at potential sources of disruption to research integrity.

Much of our oversight work is implemented through the Institute for Scientific Information (ISI), a part of Clarivate which builds on the work of Dr. Eugene Garfield – the original founder and a pioneer of information science. ISI serves as a home for analytic expertise and editorial rigour. Our global team of industry-recognised experts focus on the development of existing and new bibliometric and analytical approaches, whilst fostering collaborations with partners and academic colleagues across the global research community. In October 2020 the ISI published a global research report which speaks directly to this subject: [Research Integrity: Understanding our shared responsibility for a sustainable scholarly ecosystem.](#)⁽³⁾

Background

With the 21st century advent of cloud technology and open repositories, scientific data are now openly available. The explosion in open research – where sharing of complete datasets, rather than selected data points, is encouraged – has facilitated attempts to reproduce reported results and validate data interpretation. Without a trusted record of research, it is impossible to reliably build on previous ideas, replicate results, or effectively utilise the outcomes of research.

A [survey conducted in 2016](#)⁽⁴⁾ revealed that 70% of researchers were unable to reproduce published results, and as the Call for Evidence notes this has been termed a “reproducibility crisis,” not least for governments using taxpayer funds to pay for research and hoping to reduce costs on wasted efforts. Reproducibility can be considered part of the larger problem of research integrity, which covers issues such as data fabrication and plagiarism as well as newer forms of unethical behaviour, such as digital image manipulation and citation distortion.

We consider the driving force behind these research integrity issues to be the “publish or perish” pressure in academia and the related issue of selective reporting. Researchers are still largely hired, assessed, promoted and tenured based on the name of the journal they publish in without consideration of the broader context. Furthermore, research assessors may use the impact of the journal a study was published in as a proxy to determine the value of an individual article or researcher. Putting such emphasis on publishing in high-impact journals is detrimental to reproducibility and wider research integrity issues. Journals have an incentive to maximise their impact – typically measured through their citation performance – and are often reluctant to publish negative results or confirmatory studies as these are usually less cited than novel results. Researchers have an incentive to publish in high impact journals and may fabricate results, or over-state the significance of their results, to improve the likelihood of their manuscript being accepted for publication. The pressure to publish has also led to the birth of so-called ‘predatory journals’ that publish content without proper editorial oversight or peer review.

However, it is also important to note that not all reproducibility issues are born from intentional manipulation of the system by errant scientists – [according to Fang et al \(2012\) in a study using 2,047 retracted articles](#)⁽⁵⁾ within biomedical and life-sciences 21.3% of retractions of published research were due to human error.

The Journal Impact Factor, often referred to as the ‘JIF’ is a journal-level metric within our Journal Citation Reports. It is sometimes cited as a contributing factor to perverse incentives in academia, but our policy has always been clear - the JIF should not be used to assess individual articles or individual researchers. The founder of the ISI, Eugene Garfield, said:

“Like any other tool, the Journal Citation Reports cannot be used indiscriminately. It is a source of highly valuable information, but that information must be used within a total framework proper to the decision to be made, the hypothesis to be examined, and rarely in isolation without consideration of other factors, objective and subjective.”⁽⁶⁾

Roles and responsibilities for countering the reproducibility crisis

The responsibility for integrity in science lies with everyone involved, including researchers themselves, research institutions, funders, governments, publishers and database providers, and throughout the research process, as research integrity can be compromised at any time, from the application for funding through to final publication.⁽⁷⁾

First, researchers need to take the time to prepare, conduct and evaluate their work thoroughly, submitting their work for publication only after careful consideration and statistical validation. However, many of these factors are also influenced by their institution – for example, whether the researchers have the right resources within their labs, or whether they have struck the right balance between teaching and research time. Institutions must have clear policies around expected behavior, monitor researcher activity and take punitive action where appropriate. They must also ensure researchers are trained in all aspects of the research publishing process in order to fully support healthy research behaviour.

When it comes to assessment and directly countering the “publish or perish” culture, institutions must find **not only quantitative, but also qualitative measures to evaluate their faculty** and cease reliance on single-point metrics.⁽²⁾ This is a vital step that will tilt the scales; the pressure to produce as many articles as possible in high impact journals will be lifted and hopefully with it, the unethical practices. Institutions should consider the incentives created by evaluative frameworks carefully, and the effects they will have on researchers, many of whom have reported poor mental health as a direct result of the pressure to publish, and the scholarly record.

Standardisation and regulation of data sharing and validation is vital to research integrity and it is of funders’ interest to be the gatekeepers that safeguard it; commercial or publicly funded research which cannot be reproduced is simply a waste of money. Funders should provide institutions with affordable and accessible ways to upload, store and curate data and should take an active part in reproducibility testing. Furthermore, publishers, who already conduct screening for ethical issues such as plagiarism and conduct vital peer review should validate the results, sharing the task with the funders.

Finally, selective databases that provide content from journals and books that have passed their evaluation process need to take the final step in providing global access to trustworthy content and producing accurate and responsible bibliometric indicators and metrics, suppressing or withholding scores where possible if anomalous behavior is identified.

What policies or schemes could have a positive impact on academia’s approach to reproducible research?

A variety of policies and schemes already exist, including for example: the FAIR principles, the FORCE11 joint declaration on data citation, the Center for Open Science’s Transparency and Openness Promotion Guidelines, and guidelines from the Committee on Publication Ethics. In our experience effective policies or schemes enacted to encourage reproducible research consider the value of diversity of thought and experience, are independent from any one stakeholder, use a balance of quantitative and qualitative methods and understand the importance of partnership across the ecosystem.

The selection and evaluation of journals:

The Web of Science is the world's most trusted publisher-independent global citation database, containing more than 2 billion cited references from more than 182 million records.

To promote and protect research integrity, the Web of Science applies a [rigorous editorial selection](#) and evaluation process. For journals to be included in Web of Science, they must meet our 24 quality criteria which are designed to select for editorial rigour and publishing best practice. Application of these selection criteria screens out low quality or predatory journals so we can provide a trusted and approved list of content for search and discovery, analytics and research assessment. We do not include high citation performance as a criterion to enter Web of Science, as it is important to encourage the publication of negative results and confirmatory studies to support reproducibility. However, we do perform an additional level of evaluation, using our four impact criteria, to identify the most influential journals in their respective fields using citation activity as the primary determinant of impact.

Our [Master Journal List](#) is a free online resource listing journals that are included in Web of Science and it is used by researchers all over the world to find reputable journals for potential publication of their research. It uses the Center for Open Science's [Transparency and Openness Promotion Guidelines](#) to identify the transparent and open practices of journals.

Through support of FAIR principles and the FORCE11 joint declaration on data citation, our Data Citation Index supports reproducibility by enabling Web of Science users to locate research data sets produced as part of a piece of research through tracking citation and reuse of data sets. Data sets can be searched directly or located through a literature search where the article is linked to the indexed data sets. In all cases the data can be downloaded from the data repository to support reuse and reproducibility.

In addition, we now display open and transparent [peer review reports](#) where they are available within the Web of Science. This means that anyone can read the editorial correspondence and peer review reports relating to the final published piece of research, improving understanding of the context behind publication.

Authors:

Dr Gali Halevi, Director at the Institute for Scientific Information

Dr Nandita Quaderi, Editor-in-Chief and Vice President Editorial, Web of Science

Contact: isi@clarivate.com

References

1. Dr. Alan Finkel, Chief Scientist, Australia has publicly discussed the need for a Publication Process Quality Assurance and highlights the selection and evaluation criteria in the Web of Science Core Collection as a potential framework which satisfies this need.
<https://www.chiefscientist.gov.au/2019/09/14243>
2. Adams J, McVeigh M, Pendlebury D A, Szomszor M. [Profiles, not metrics](#). Institute for Scientific Information, (2019)
3. Szomszor, M. and Quaderi, N. (2020) [Research Integrity: Understanding our shared responsibility for a sustainable scholarly ecosystem.](#), <https://clarivate.com/webofsciencegroup/wp-content/uploads/sites/2/2020/10/ISI-Research-Integrity-Report.pdf>
4. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016).
<https://doi.org/10.1038/533452a>

5. Ferric C. Fang, R. Grant Steen, and Arturo Casadevall, Misconduct accounts for the majority of retracted scientific publications, PNAS October 16, 2012 109 (42) 17028-17033; <https://doi.org/10.1073/pnas.1212247109>
6. Garfield, E. (1975) "Preface and Introduction to Journal Citation Reports –Vol. 9 of the SCI, 1975."
7. Sivasubramaniam, S.D., Cosentino, M., Ribeiro, L. *et al.* Unethical practices within medical research and publication – An exploratory study. *Int J Educ Integr* **17**, 7 (2021). <https://doi.org/10.1007/s40979-021-00072-y>

(September 2021)