Paying it forward – publishing your research reproducibly

It might be fair to say we have entered the era of irreproducible science, write Martijn Roelandse and Anita Bandrowski

●●●

In a recent study, it was estimated that 50 per cent of the US preclinical research spend was not reproducible. That is a total of 28 billion USD or the equivalent of 600,000 annual postdoc salaries! On top of that, the success rates for new development projects in Phase II trials have fallen from 28 per cent to 18 per cent in recent years.

Francis Collins, director of the National Institute of Health, stated: 'A growing chorus of concern, from scientists and laypeople, contends that the complex system for ensuring the reproducibility of biomedical research is failing and is in need of restructuring.'

The majority of issues around irreproducibility are flaws in reference material associated with the unreliable identification of source materials used in the preclinical study, particularly contaminated, mishandled, or mislabeled biological reagents like antibodies or cell lines.

One of these flaws, unreliable identification of materials, was noted recently in an editorial expression of concern issued for an article in *Science*. The crux of that problem seemed to be that one of the authors picked a virus strain that caught the other authors by surprise; an error that was not caught. Three studies and a clinical trial attempted to replicate the findings, without success, which dealt a promising HIV cure the final blow.

Another flaw causing hundreds of scientists to create an organisation devoted to cell line authenticity is the use of problematic cell lines. For example, more than 300 studies had used a breast adenocarcinoma cell line before it was found to be derived from human ovarian carcinoma cells. $100 million of research funding may have been spent using this misidentified cell line alone.

At this stage, we will not recount here how the more than 1,000 cell lines, in which problems were reported, continue to contaminate the cancer literature (Freedman et al, 2015). Instead, I would like to draw your attention to a simple act that may substantially reduce the use of problematic cell lines. A recent study by Babic et al. (2019) showed that in papers that identify cell lines through RRIDs, research resource identifiers, the use of problematic cell lines was substantially lower than in those that did not. RRIDs were introduced in 2014 and since then have caught the attention of many publishers such as Cell Press, because it is a fairly simple method for disseminating important information about reagent quality before the paper is published, saving the need to issue editorial expressions of concern or even retractions.

In fact, later that year a group of editors representing more than 30 major journals, representatives from funding agencies, and scientific leaders drew up a list of Principles and Guidelines for Reporting Preclinical Research. These identified four key areas:
• Scientific rigour (or rigorous experimental design);
• Scientific premise (or strength of the key data supporting the proposed research);
• Identification of key resources; and
• Sex and other biological variables.

"Scientists have certainly become aware of the problems with reproducible research"

Since then several projects have been initiated, with varying success.

Scientists have certainly become aware of the problems with reproducible research, as evidenced by a survey conducted by Nature. Some journals implemented checklists that address many of the principles for reporting of preclinical research, and the effect on the top journal, *Nature*, has been positive with authors making explicit aspects of their methods such as whether or not they blinded any aspect of their study to reduce investigator bias. Indeed, last year, a group of publishers have even taken the important step of creating a multi-publisher checklist so wherever authors decide to publish, the standard would be the same (See MDAR project).

In the meantime, more and more manuscripts are being submitted per year and the pressure on journals and reviewers, to assess the quality of the work is increasing. So what can possibly be done to maintain or improve the quality of peer review?

One answer is to pay for peer review. There certainly are scientists that would like to make a little extra money on the side, and well-resourced journals such as *eLife* routinely use professional review as part of their process. This eliminates poor quality and inconsistent manuscripts from being sent to traditional peer reviewers, streamlining the process.

Another answer is to do away with peer review altogether. Indeed, preprint servers are being used more as sources of information, and these manuscripts are now frequently cited in pre-peer review work as well as peer-reviewed work.

Perhaps a more interesting solution, at least from the technology perspective, is to offload part of the peer review process onto machines. A recent survey on twitter by Helen King resulted in a plethora of tools that in some way or form support the publication process. Nearly 50 percent of those play a role in the submission process, performing technical checks, editorial support, metadata extraction or language polishing. A portion of the technical checks is around reproducibility where Barzooka, OddPub, JetFighter, limitations-finder, seek&blastn, Ripeta, and SciScore may lend support.

So what are these tools and what can they tell us about the manuscript?
JetFighter checks for colour-blindness compliance of images;
Barzooka finds bar graphs and attempts to figure out if the authors are using these for continuous variables (a bad way to represent continuous data);
OddPub checks for statements about open code and open data;
Limitations-finder pulls out authors

sentences describing how their study was limited;

Seek & Blastn checks for common oligonucleotide problems;

Ripeta checks several rigour criteria such as whether blinding was used in the study;

SciScore.com checks the rigour and also checks reagents such as antibodies, cell lines, and organisms as well as plasmids, and software tools.

These tools provide output, which can alert authors to a particular problem with their manuscript. However, these tools can also alert the reviewer or editor whether there is a particular problem or problems with the manuscript at hand. Several of these tools have been deployed across different parts of the literature. For example, JetFighter pings authors of any preprint in which a colour-blind non-compliant figure is detected. We understand that some authors do not appreciate this, but it is difficult to know if there is a better way to raise awareness.

Several of the other tools (SciScore, Limitations-finder, OddPub, Barzooka) have recently been deployed together and provide a single report for COVID literature (see example report). It is still too early to tell whether there will be any improvement in the literature based on this work, but given the pace of publishing, especially in the preprints, a set of tools seems like an important step towards better science.

Indeed, the editors of AACR, have taken a long look at reproducibility issues in cancer biology as described last year in

"These tools provide output, which can alert authors to a particular problem with their manuscript"

an editorial (Dang, 2018), however, theory must be put into practice. Although RRIDs have been encouraged in the instructions to authors of AACR journals over the last year, these recommendations have not been followed widely by all authors, most likely because it is a tedious task to do. Moving forwards, this recommendation will become a little more stringent as we have deployed SciScore (the Rigour criteria and RRID checker) in eJournalPress for all AACR journals as part of the submission workflow. All authors submitting a manuscript to AACR will receive a rigour adherence and key resource report that will help them to improve their manuscript and, if followed, should make the AACR's research more reproducible. **Ri**

Martijn Roelandse is founder/consultant at martijnroelandse.dev

Anita Bandrowski runs the RRID initiative and is the creator of SciScore..

A fully-referenced version of this article is available at www.researchinformation.info

Glossary

Rigour is defined as 'the strict application of the scientific method to ensure robust and unbiased experimental design' (National Institutes of Health, 2018).

RRIDs (research resource identifiers) are ID numbers assigned to help researchers cite key resources (antibodies, model organisms and software projects) in the biomedical literature to improve the transparency of research methods. RRIDs are also provided for other resources such as organisms, plasmids, and software tools. When RRIDs are not present in the RRID portal, authors are encouraged to register a new one.

MDAR (materials design analysis reporting) framework is a checklist for authors to use, and an elaboration document with background and instructions. The project components contain data from author and editor surveys and coder data from the evaluated checklists.

SciScore is an advanced text mining based tool that checks the methods section for the use of RRIDs and for compliance with the NIH rigour and transparency criteria. The latter include proper authentication of cell lines, an important step for many cancer researchers.

Automated Screening Working Group is a group of software engineers and biologists passionate about improving scientific manuscripts at scale. Our goal is to process every manuscript in the biomedical sciences as it is being submitted for publication to improve that manuscript. Each tool checks for a different set of transparency criteria, but together we can shed light on what your manuscript. We will build pathways for the tools to work together.

Product Spotlight

SciScore

The right time to do the right thing about scientific reproducibility.

One thing that this pandemic has taught us is that not everything we do must be done in the same way as it "always was". Science is now happening at a furious pace. Journals have been inundated by "that other paper" that has just been sitting around in a file drawer, until that is, the fast pace of life suddenly stops. So here we are, publishing so fast that it exasperates any previous problems with quality checks. Worse still, scientists working on COVID-19 argue that their studies must be released at the earliest.

So what is a publisher (traditional or preprint) to do? How can we possibly maintain quality in this type of environment?

Clearly we need a reviewer that can read at a thousand words per second, never gets tired, and does not refuse to review the paper.

That reviewer actually exists - meet SciScore. com. It can read a methods section of a manuscript in about a minute, and provides a report to editors, reviewers or authors about common rigor criteria that have or have not been addressed. SciScore has already read and reviewed all of the COVID manuscripts, and posted the results to twitter and Hypothes.is. SciScore is also 'reading' all AACR papers giving the hard-working editors a hand, and will soon be working for any journal that works with the Editorial Manager platform.

Interesting for your journal?
Check sciscore.com/product

ALPSP Awards for Innovation in Publishing 2020

Finalist