Recent developments in scholarly publishing: a view from the life sciences

Stephen J Eglen (orcid), Ross Mounce (orcid), Laurent Gatto (orcid), Adrian M Currie (orcid).

Introduction

In some ways, scholarly publishing has not changed much in the last ten years. Publishing in prestigious top-tier journals is still perceived as critical for career progression (especially gaining promotion and grants). Likewise, journal metrics continue to dominate in the evaluation of a paper's research, rather than the paper's contents (Brenner 1995). Against the backdrop of highly competitive job and grant markets, factors such as these encourage narrow research agendas and tie researchers (particularly in early career) to placing work in exploitative publishers who draw significant funds from academic work. Further, standard publishing criteria, especially for instance on publishing statistically significant, positive results, creates biases across published studies. However, there are several reasons for optimism that the nature of scientific publishing will improve. Here we outline some recent developments in the life sciences.

Preprints

Since 1991 ArXiv (Ginsparg 2017) has become a standard tool for physicists to rapidly disseminate their research findings. On the surface, it does not provide much beyond a collection of PDFs grouped by topic. In certain fields, being the first to publish on the ArXiv is considered to be the key step, although subsequent journal publication is still the norm. Although ArXiv hosts papers in quantitative biology, it was initially assumed that biologists would not adopt a preprint culture: publishing a preprint might prevent subsequent publication in a top-tier journal, or leading to scooping by another group. There is some historic justification of these concerns, as an NIH experiment in preprints was effectively halted in the 1960s by journals' refusal to accept preprints for submission (Cobb 2017).

BiorXiv, launched in 2013, has overcome these concerns. Researchers in diverse areas as ecology, neuroscience and genomics are uploading preprints and choosing to share their work ahead of publication. There are many reasons for this usage:

- Sharing work before submitting to a journal allows for community feedback.
- Sharing work at the time of submission means that the community can read the work months (or years) before the work eventually appears in print.
- Journal editors are browsing the bioRxiv and soliciting that relevant papers be submitted to their journal. (How widespread this is, I'm not sure.)
- BioRxiv preprints can be transferred rapidly to journal submission systems rather than going through (often lengthy) direct submission to the journal.
- Several funding agencies, including NIH and UKRI, allows preprints to be listed and cited in grant applications.

Several other preprint servers have been created in recent year, in particular PeerJ Preprints, OSF Preprint servers, and preprints.org, although to date BioRxiv is the dominant repository. Unlike a few years ago, most journals in the life sciences no longer see prior appearance in a repository as a block to formal publication.

Overcoming the reproducibility crisis

According to recent surveys, life scientists across many domains believe there is a "reproducibility crisis" in science: i.e. many key key findings in publications are either not independently verified,

or fail verification when it is done (Baker 2016). The traditional publishing system must take some responsibility for these low-levels of reproducibility. However, here we list three encouraging developments that should promote reproducibility.

Preregistration papers. An encouraging response to this crisis has been the adoption of preregistration papers (Nosek et al. 2018). These papers typically describe the introduction and methods sections of a paper, and are peer-reviewed *before* the study is actually performed. This allows reviewers to improve the study design and commits researchers to hypotheses that they wish to study along with their statistical analysis. Once the pre-registration study is approved, it is then published. After the research is completed, another paper describes the results of the study using the pre-registered methods. (Additional findings can be reported, but are clearly marked as such.) Preregistration is most prevalent today in psychology; The Center for Open Science Preregistration Challenge https://cos.io/prereg/ is helping to popularise this notion more broadly.

Stronger data sharing policies and community expectations. Both funders and journals are now making stronger statements about what research materials (data, computer programs, reagents) should be shared upon publication of the corresponding articles. Although these policies should increase data availability and reuse, the current compliance rates are quite low (Federer et al. 2018). Given that it can take considerable time and effort (for both researchers and journals) to ensure data is appropriately shared, these low-uptake rates are perhaps expected. To reward authors for this work, "data papers" (a paper that simply describes the data) are becoming more prominent, e.g. in journals like *Scientific Data* and *Gigascience*.

Reproducible manuscripts are documents that contain the main text as well as the code to generate tables, figures and results have been around for decades, and have been widely used in many research fields (Buckheit and Donoho 1995). However, even though researchers have been committed to reproducible research, the reproducibility of the final outputs were generally broken upon submission to journals. Researchers have released reproducible versions of their work in parallel to the journal articles (Gatto and Christoforou 2014; Breckels et al. 2018). Recently, some journals have moved closer to publishing reproducible manuscripts, by working towards a reproducible documement stack, or supporting reproducible figures.

Other recent innovations of note

ORCiD https://orcid.org provides a persistent, unique digital identifier for researchers. Many journals now require that at least one author verifies their identity as author using ORCiD [https://orcid.org/content/requiring-orcid-publication-workflows-open-letter].

DORA https://sfdora.org/ is a declaration for individuals and institutions to commit to evaluating research based on its content rather than metrics. Most UK funders have signed, although only a few universities have signed. See also the Leiden Manifesto for Research Metrics (http://www.leidenmanifesto.org/).

Published Peer Review Reports. Many journals now already or have pledged to provide greater transparency about the quality of peer review they provide by publishing the content of the reviewer reports alongside published articles. Notably two large open access publishers PLOS and MDPI are amongst those that are pledging to provide greater transparency from 2019 http://asapbio.org/letter. PubPeer https://pubpeer.com/ allows reviewers to 'claim' metadata records on their profile for peer reviewing and editorial work they have done.

Post publication peer review. A journal may immediately publish a paper upon submission; reviews are then sought for the preprint and made public. If sufficient reviewers support publication, the article is formally accepted and e.g. listed on Pubmed. Leading examples of this approach are F1000 Research, who provide the infrastructure for several institution- and funder-specific journals, such as *Wellcome Open Research* and *Gates Open Research*.

Format free submissions. Journals have traditionally imposed strict formatting requirements for manuscripts before peer review. As editors at top-tier journals 'desk reject' most submis-

sions before peer review, this leads to many wasted hours (Budd 2017). Gradually life science journals are now dropping these formatting requirements for initial submissions, instead allowing "format free" submissions [Khan, Montenegro-Montero, and Mathelier (2018); see also https://asntech.github.io/format-free-journals/].

Funder mandates and compliance

Key funders in the UK have had policies in place supporting open access for many years. In particular, the Wellcome Trust has mandated Open Access for publications funded by them since 2006, with sanctions for non-compliance. Compliance rates (around 90%) are highest for the WT, as of October 2017 (Larivière and Sugimoto 2018), with compliance for other main funders varying at 70-90%. Where work has been supported by relevant funding agencies, our experience to date is that funds have always been available to support APCs. However, one of us [SJE] has experienced difficulties in finding APCs for papers summarising work supported by internal, rather than external, funds.

One perhaps unintended consequence of these policies has been that most traditional journals have established a "hybrid" model of publishing, with APCs that on average exceed those in pure OA journals (Pinfield, Salter, and Bath 2017). This hybrid model of publishing has shown little signs to date of disappearing, as e.g. funds from Wellcome Trust have supported high APCs. The success of OA publishing however has meant that government-provided funds can often no longer cover all APCs and UK institutions are beginning to restrict the choice of journals for which APCs will be paid. However, The OA publishing world is due to change dramatically in 2020 with the recent announcement of "plan S" (Schiltz 2018), a European initiative to enforce OA, cap APCs and prohibit publishing in hybrid journals. Whilst we support the notions underlying plan S, its success will depend on further implementation details (e.g. the nature of the APC cap, recognition of green and diamond OA).

Concluding remarks

Current publication practices can often lead early career researchers to be 'Bullied into Bad Science' (http://bulliedintobadscience.org/). We have outlined several recent developments that we hope present alternatives to the traditional hierarchy of scholarly publishing. These developments should help reduce the pressure on early career researchers that they currently face in the "publish or perish" culture. We encourage the adoption of the above open practices to help create a more ethical research environment.

Glossary/abbreviations

Perhaps need a list of key abbreviations/terms that are jargon (APCs/hybrid/diamond OA).

References

Baker, Monya. 2016. "1,500 Scientists Lift the Lid on Reproducibility." *Nature* 533 (7604): 452–54. doi:10.1038/533452a.

Breckels, Lisa M, Claire M Mulvey, Kathryn S Lilley, and Laurent Gatto. 2018. "A Bioconductor Workflow for Processing and Analysing Spatial Proteomics Data." *F1000Res.* 5 (July).

doi:10.12688/f1000research.10411.2.

Brenner, Sydney. 1995. "Loose End." *Curr. Biol.* 5 (5): 568. doi:10.1016/S0960-9822(95)00109-X.

Buckheit, Jonathan B, and David L Donoho. 1995. "WaveLab and Reproducible Research." In *Wavelets and Statistics*, 55–81. Lecture Notes in Statistics. Springer New York. doi:10.1007/978-1-4612-2544-7_5.

Budd, Julian. 2017. "Publishing: Reformatting Wastes Public Funds." *Nature* 543 (7643): 40. doi:10.1038/543040e.

Cobb, Matthew. 2017. "The Prehistory of Biology Preprints: A Forgotten Experiment from the 1960s." *PLoS Biol.* 15 (11): e2003995. doi:10.1371/journal.pbio.2003995.

Federer, Lisa M, Christopher W Belter, Douglas J Joubert, Alicia Livinski, Ya-Ling Lu, Lissa N Snyders, and Holly Thompson. 2018. "Data Sharing in PLOS ONE: An Analysis of Data Availability Statements." *PLoS One* 13 (5): e0194768. doi:10.1371/journal.pone.0194768.

Gatto, Laurent, and Andy Christoforou. 2014. "Using R and Bioconductor for Proteomics Data Analysis." *Biochim. Biophys. Acta* 1844 (1 Pt A): 42–51. doi:10.1016/j.bbapap.2013.04.032.

Ginsparg, P. 2017. "Preprint déjà Vu: An FAQ." *ArXiv [Cs.DL]*, June.

Khan, Aziz, Alejandro Montenegro-Montero, and Anthony Mathelier. 2018. "Put Science First and Formatting Later." *EMBO Rep.* 19 (5). doi:10.15252/embr.201845731.

Larivière, Vincent, and Cassidy R Sugimoto. 2018. "Do Authors Comply When Funders Enforce Open Access to Research?" *Nature* 562 (7728). Nature Publishing Group: 483. doi:10.1038/d41586-018-07101-w.

Nosek, Brian A, Charles R Ebersole, Alexander C DeHaven, and David T Mellor. 2018. "The Preregistration Revolution." *Proc. Natl. Acad. Sci. U. S. A.* 115 (11): 2600–2606. doi:10.1073/pnas.1708274114.

Pinfield, Stephen, Jennifer Salter, and Peter A Bath. 2017. "A 'Gold-Centric' Implementation of Open Access: Hybrid Journals, the 'Total Cost of Publication,' and Policy Development in the UK and Beyond." *Journal of the Association for Information Science and Technology* 68 (9): 2248–63. doi:10.1002/asi.23742.

Schiltz, Marc. 2018. "Science Without Publication Paywalls: cOAlition S for the Realisation of Full and Immediate Open Access." *PLoS Biol.* 16 (9): e3000031. doi:10.1371/journal.pbio.3000031.