

Preprint servers: a “rush to publish” or “just in time delivery” for science?

Alan R Smyth, Professor of Child Health, University of Nottingham.

Claire Rawlinson, Publisher at BMJ and co-founder of medRxiv.

Gisli Jenkins, NIHR Research Professor, University of Nottingham.

At the time of writing, the World Health Organisation has declared COVID-19 (novel corona virus) a pandemic and health-care systems around the world face the biggest public health challenge for a generation. As never before, the clinical and scientific communities, governments and the public need access to robust data to facilitate evidence-based decision making. The foundation of a reliable research literature is effective peer review. However, in the last decade, there has been debate over the effectiveness of peer review. Research in the field has shown that prestigious journals reject manuscripts which go on to be highly cited elsewhere¹. The peer review process introduces considerable delay in making research findings publicly available² and yet there is no correlation between the number of rounds of manuscript review and revision and the subsequent citation count for the paper¹.

The stage is therefore set for the entrance of a new actor – the preprint server. A recent article (posted as a preprint) identified almost 50 preprint servers, open research bundles or other disruptive technologies which have entered the field in recent years³. The distinctive feature of preprint servers is that they do not undertake peer review but restrict scrutiny to basic screening and legal checks, such as for plagiarism. However, preprint servers do allow online comments from the scientific community – a form of “crowd-sourced” peer review. Substantive comments can

allow the amendment of the manuscript before subsequent submission to a conventional peer-reviewed journal.

One of the first such preprint servers was arXiv (founded in 1991) to serve the physics, astronomy and mathematics research communities. The established player in the life sciences field is bioRxiv, founded in 2013 and currently hosting almost 80,000 articles. The new kid on the block serves medical sciences: medRxiv was launched in June 2019⁴ and is currently hosting more than 2,500 articles (including over 850 on COVID-19).

At *Thorax* we embrace this new pathway to publishing medical research findings and we welcome the submission of manuscripts which have previously appeared on a preprint server. We do, however, ask all submitting authors to make this clear in the covering letter at the time of submission. The first batch of 10 articles, which previously appeared as preprints, have been through peer review with *Thorax*. The acceptance of articles which have previously appeared as a preprint is now widespread amongst medical journals^{5,6}. Acceptance of preprints is, however, not universal and authors are well advised to check the guidelines of their target journals before they post a preprint.

So do preprint servers live up to the hype? Do they allow “just in time delivery” for science? Are they an innovation whose hour has come during a public health crisis such as the COVID-19 pandemic? In medicine (as distinct from the life sciences) there are some declared limitations. MedRxiv carries a disclaimer on its home page which says that preprints: “...*should not be relied on to guide clinical practice or*

health-related behaviour and should not be reported in news media as established information.” The editor of one leading medical journal has previously suggested, with reference to preprints, that the “...*rush to publication*...” is an “... *editorial and scientific mistake*...”⁷. This assertion is supported by the fate of a recent preprint, suggesting similarities between the novel coronavirus and HIV virions⁸. This preprint was withdrawn, after receiving over a hundred (mostly critical) online comments – although it could be argued that this represents effective, informal peer-review. Retraction of a misleading article, published in a conventional journal, would have taken a lot longer. However, some data suggest that there is little time for this informal peer review. Some journals are publishing articles only a few months after their appearance on bioRxiv⁹ – suggesting that journal submission occurs shortly after preprint posting. This would leave little time for feedback on the preprint to allow amendment or correction prior to journal submission.

It has also been suggested that preprints have helped to drive the early discourse and have influenced policy-making at the start of the COVID-19 pandemic¹⁰. The UK government cited a preprint in their first coronavirus action plan¹¹. There has been an exponential rise in preprint articles deposited during the pandemic¹² reflecting the urgent need and desire for the latest information. In the face of a pandemic such as COVID-19, “just in time delivery” is critical for research. Indeed, preprint servers allowed the rapid publication of data used to allow vital modelling the trajectory of the pandemic¹³.

Both traditional publishers and preprint servers are adapting to the demand for rapid access to data on COVID-19. Many traditional journals are fast tracking COVID-19

publications with over 300 published already this year¹² and medRxiv has been more sympathetic to studies with a small number of participants because clinical data have not been available in any other form.

In due course, when the COVID-19 curve (flattened or otherwise) hits baseline, researchers and journals must use the preprint literature wisely and as it is intended – as a way to share research data rapidly before formal expert review in a journal. Any individual claims should be treated with healthy scepticism, until verified by peer review. Such scepticism can be difficult in the current frenzied environment.

Therefore, it is vital that the research community does not misuse the benefits of preprints to short-circuit the quality controls that drive peer reviewed publication. Similarly, it is crucial that journals streamline, but maintain high-quality peer review processes, and in times of crisis make content free to access. At Thorax we are committed to supporting preprints whilst continuing to provide robust, rapid and fair peer review to maintain confidence in the quality of the published work.

REFERENCES

1. Siler K, Lee K, Bero L. Measuring the effectiveness of scientific gatekeeping. *Proc Natl Acad Sci USA*. 2015;112(2):360-365. 10.1073/pnas.1418218112
2. Fernandez-Llimos F, & Pharmacy Practice peer reviewers. Peer review and publication delay. *Pharm Pract (Granada)*. 2019;17(1):1502-1502. 10.18549/PharmPract.2019.1.1502
3. Penfold N, Murphy F, Kirkham J. Practices and policies of preprint platforms for life and biomedical sciences. *Zenodo*. 2020. <http://doi.org/10.5281/zenodo.3700874>
4. Rawlinson C, Bloom T. New preprint server for medical research. *BMJ*. 2019;365:l2301. 10.1136/bmj.l2301
5. Verma IM. Preprint servers facilitate scientific discourse. *Proc Natl Acad Sci USA*. 2017;114(48):12630. 10.1073/pnas.1716857114
6. Bell SC, Flume PA, Castellani C. Seven P's of publication practices. *J Cyst Fibros*. 2020 (epub ahead of print). <https://doi.org/10.1016/j.jcf.2020.02.007>

7. Bauchner H. The rush to publication: An editorial and scientific mistake. *JAMA*. 2017;318(12):1109-1110. 10.1001/jama.2017.11816
8. Pradhan P, Pandey AK, Mishra A, et al. Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag. *bioRxiv*. 2020:2020.2001.2030.927871. 10.1101/2020.01.30.927871
9. Abdill RJ, Blekman R. Tracking the popularity and outcomes of all bioRxiv preprints. *Elife*. 2019;8. 10.7554/eLife.45133
10. Majumder M, Mandl KD. Early in the Epidemic: Impact of Preprints on Global Discourse of 2019-nCoV Transmissibility. *SSRN*. 2020. <http://dx.doi.org/10.2139/ssrn.3536663>
11. UK Department of Health & Social care. Coronavirus: action plan, A guide to what you can expect across the UK. <https://www.gov.uk/government/publications/coronavirus-action-plan/coronavirus-action-plan-a-guide-to-what-you-can-expect-across-the-uk>. Accessed 3 March, 2020.
12. Callaway E, Cyranoski D, Mallapaty S, Stoye E, Tollefson J. The coronavirus pandemic in five powerful charts. *Nature*. 2020;579(7800):482-483. 10.1038/d41586-020-00758-2
13. Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Dis*. 2020. 10.1016/S1473-3099(20)30144-4