

Pondering Preprints and Progress

Tracey A DePellegrin

Preprints have long been established in some fields, and they are on the rise in many others. It's easy to see why. Scientists painstakingly conduct research for years, and the primary output of this work is an article published in a scholarly journal. Sometimes that process takes a long time, or at least longer than researchers—and their audiences—want to wait to communicate findings.

Researchers in the life sciences are increasingly turning to preprint servers like bioRxiv (<https://www.biorxiv.org/>), which allows authors to deposit unpublished life science manuscripts, and posts those manuscripts online as a “preprint”. Preprints are often simultaneously submitted to journals and deposited on a preprint server. Anyone with an internet connection can read and comment on the paper through the preprint server before it is published.

Evolving quickly are the ways journal publishers handle manuscript submissions that have been posted as preprints. Just 5 years ago many journals prohibited the submission of manuscripts that had been posted as preprints. Today most journals welcome preprints. Indeed, some editors scour preprint servers to recruit submissions.

The adoption of preprints varies by field. As of April 2018, bioRxiv has preprinted over 23 500 manuscripts from 134 000 authors representing 8 500 institutions from over 100 countries.

The arXiv preprint server (<https://arxiv.org/>), a mainstay in the physics, math, and computer science communities since 1991, hosts over 1.3 million preprints. Other preprint platforms have been established more recently, including *PeerJ Preprint* (<https://peerj.com/preprints-search/>), and field-specific preprint servers such as ChemRxiv (<https://chemrxiv.org/>), EarthArXiv (<https://eartharxiv.org/>), engrXiv (<https://engrxiv.org/>), and ESSOAr (<https://www.essoar.org/>).

Questions remain as preprints are becoming part of the scholarly publishing ecosystem. Can preprints be cited? Should editors use preprints and their posted comments as part of the manuscript peer review? How can journals ensure that only one version of record exists? Additional questions arise when one thinks of preprints in the field of medicine and health sciences, and how these sometimes life-and-death scientific topics might be used. MedRxiv (<http://yoda.yale.edu/medrxiv>), maintained by the Yale Open Data Access Project (affectionately named YODA), was created

with those types of preprints in mind and has myriad policies around data access geared toward a niched audience.

Preprints are rife with opportunity. Comments about preprints arising on social media or a preprint server can be useful, in particular as a way for authors to garner initial feedback. Scientists can also establish precedence of ideas and connect their work to readers as soon as possible. Journal clubs discussing preprints are popping up, providing early career researchers with ways to engage with one another and with research. Some journals have “preprint editors” who trawl the servers in the hopes of recruiting manuscripts for their journals. Scholarly societies and other organizations can formalize preprint reviews, and when coupled with robust peer review in a journal, preprints can provide authors and readers with the best of all worlds. Scientists and readers can enjoy all the benefits of posting their preprint while it's undergoing review, thereby accelerating access to the work and still realizing the wins from a polished, revised paper that was peer reviewed, edited, published, and promoted (not to mention 100 other steps!).

Large, well-funded labs are flocking to preprints. Indeed, those labs are some of the most poised to submit to preprint servers. They have myriad colleagues able to read and revise (before submitting), and some have communications managers. However, smaller labs with fewer resources may benefit most from the structure of journal peer review, editing, and article amplification and promotion.

There are many benefits to preprints covered in already-published articles, so rather than delve too deeply here, I have included a list for further reading, below. Full disclosure: GSA Journals were the first to partner with bioRxiv to allow submissions at GENETICS and G3 to be seamlessly transferred to bioRxiv. See our editorial at <http://asapbio.org/synergy> and <http://genestogenomes.org/gsa-journals-partner-biorxiv/>. This arrangement with bioRxiv has worked for us since 2014.

As with most innovations, preprints come with some drawbacks, most of which I suspect will be smoothed out over the coming years. People worry about being scooped. Whether this fear is based on facts is unclear. Some scoff (one of my favorite bits of “anecdota” involves the loud vocalization that this fear is unsubstantiated because they haven't heard of this happening), but in today's hypercompetitive atmosphere, it's hard to blame scientists who have misgivings. Sure, the papers are free from those pesky editor and reviewer requests for additional

experiments—but what if those additional experiments are actually necessary to support the paper’s conclusions? If preprints are posted but never published in a journal, will they suffer from a lack of tagging, indexing, readership, promotion, and archiving (and 100+ other things journals do—see the post from Anderson, listed in the Further Reading, below)? With no vetting of content, it’s not clear if a bogus preprint that contains misinformation on health, therapeutics, or biosecurity will mislead the public or the press. Some preprint servers guard against that kind of thing, but are all taking on that responsibility?

Related is the nature of preprint servers—low-cost to run, low activation/energy to submit—means that editorial offices and editors aren’t combing through the submission and evaluating for the quality, the presence of data, an indication that all authors agreed to the submission (and how it will be used), and the presence of markers of scientific integrity.

Data sharing is one area preprints may lag behind. In an ideal world, authors are generous with data sharing and provide the raw data to support the paper. For myriad reasons not all scientists, however, are this open about data sharing (e.g., in genetics, work with populations associated with proprietary companies like animal breeders). What’s a paper without data to back it up? Many journals require raw data before a manuscript will be considered for submission. Authors who are unable or unwilling to provide this data may not publish their papers in these journals.

My sense is that preprints aren’t supplanting journals (not yet anyway). The two can co-exist peacefully and productively, and serve to improve the ecosystem and the scientist experience. We ought to understand the value of preprints and what proponents are saying, as well as the potential drawbacks.

As editors and publishers, we are entrusted by authors with years of their hard work. We must continue to carry out peer review and innovate process and policy such that we not only uphold, but also promote its integrity. We provide checks on ethics, and ensure data availability and quality. We make sure that papers are properly tagged, indexed, discoverable, readable, and citable. We highlight, promote, and discuss not just the science, but the stories and people behind the discoveries. We help improve a paper’s impact not just for today, but for years to come.

I call on each of you, as members of the scholarly publishing community, to reaffirm your role as author advocates. Ideally, providing authors with robust, ethical, timely peer review of manuscripts and thoughtful decision letters from editors should improve their papers and their science. I think we must pay attention to what the market demands; if our communities want to use preprint servers, we owe it to them to understand

how preprints might complement what we offer. We must work hard to serve our authors to address their changing needs and to provide our readers with articles worth their valuable time, or risk being left behind.

This issue of *Science Editor* features some must-reads, including an insightful piece about author surveys by Jessica Rucker and Jody Plank, PhD, from the American Chemical Society. They describe a Herculean effort to conduct a longitudinal study using rolling author surveys starting in 2015. To date, they have amassed more than 34000 responses, and an almost-equal number of open-ended responses. This candid article proves interesting and useful, providing tidbits about their findings and lessons learned about understanding and engaging authors.

In “Reevaluating the Quest for Negative Results,” Lenny Teytelman, PhD, asks readers to challenge some of our assumptions about publishing these papers. He asserts that there is no shortage of venues for publishing negative results; rather, there is a paucity of submissions. What follows is a discussion of the value of negative results and the nuances of understanding and publishing them.

Lorinc *et al.* in their Perspectives article “The Painful Publishing Process: A Request to Simplify Bureaucratic Requirements” make a plea for editors to allow manuscript submissions in myriad formats. It’s a case you’ll want to hear out.

We also introduce a column, “Fire of the Week,” written by Emilie Gunn, who has also agreed to act as Editor for the regular feature. These pieces will describe the types of events (if not the actual ones) most of us can relate to—urgent, adrenaline-fueled emergencies that crop up, if not hourly or daily, than most certainly weekly. Emilie dissects what happened and why, plus the steps she took to mitigate the problem. These “fires” are so common that we are hoping you will share with *Science Editor* stories about your own and how you put them out.

Further Reading

- <http://www.sciencemag.org/news/2017/09/are-preprints-future-biology-survival-guide-scientists>
- <http://mbio.asm.org/content/9/2/e00516-18.long>
- <http://blogs.plos.org/plos/2018/05/power-to-the-preprint/>
- <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005473>
- <http://asapbio.org/about-2>
- <https://scholarlykitchen.sspnet.org/2018/02/06/focusing-value-102-things-journal-publishers-2018-update/>