Preprints, postprints, and VORs

So, you've met the OA colour scheme by now, and hopefully that was simple enough. Now, we're going to add another dimension to this based on the *types* of articles represented.

Some useful definitions to start things off here include:

- **Preprint** - Version of a research paper, typically prior to peer review and publication in a journal.
- **Postprint** - Version of a research paper subsequent to peer review (and acceptance), but before any type-setting or copy-editing by the publisher. Also sometimes called a 'peer reviewed accepted manuscript'.
- **Version of record (VOR)** - The final published version of a scholarly research paper, after undergoing formatting (and any other additions) by the publisher.
- **e-Print** - Version of a research paper posted on a public server, independently of its status regarding peer-review, publication in print, etc. Preprints, postprints and VORs are forms of e-Prints.

[INSERT 04_open_access_archives.png here; credit: Patrick Hochstenbach, CC BY]

- **Accepted author manuscript (AAM)** - The version of a manuscript that has been accepted by a publisher for publication. Can often be a postprint.

It really helps to think of these as parts in a publication pipeline. Preprints are pre-peer review; postprints/AAMs are post-peer review; VORs are post-production.

The case for preprints

Now, none of these concepts are particularly new. The first experiments with preprints go all the way back to the 1960s (Till, 2001). Starting with the biomedical sciences, this was before the internet so part of a paper-based preprint culture. The first free scientific online archive was arXiv, started in 1991, initially a preprint service for physicists, initiated by Paul Ginsparg. Since its origins, self-archiving has now become the norm in many sub-fields of physics, especially high-energy physics. Now, arXiv includes papers from related disciplines including computer science, mathematics, nonlinear sciences, quantitative biology, quantitative finance, and statistics. arXiv now also includes postprints as well as preprints. Some would argue that one of the reasons the World Wide Web even exists is for preprints. Now, virtually all research disciplines have their own dedicated preprint servers, with more being established on a frequent basis.

At the present though, the perception of preprints is highly contingent on the history of their use in disciplines. For example, there are big differences between physics and the life sciences community, as illustrated by Neylon et al.., 2017 below. This makes perfect sense - science is simply conducted differently between disciplines, and peer review is often held in a different regard too.

[INSERT NEYLON_preprints here]

The main power of preprints is that they enable authors to make their manuscripts freely and publicly available in parallel to, or before, submitting them to journals for formal peer review. As such, they help to communicate research results much faster than traditional methods - see Desjardins-Proulx et al., 2013 and Sarabipour et al., 2019.

Other advantages include:

- Rapid dissemination to a wider audience;
- Immediate visibility of research and facilitate networking;
- Demonstration of research progress, particularly for early-career researchers;
- Availability for wider feedback/review from the research community;
- Establishment of precedence or intellectual priority for research discoveries;
- Potential citation advantage due to earlier and wider availability of research;

- Accelerating training time and optimising research design and quality;
- Publishing Open Access even with limited or no funds;
- Commenting on preprints for developing peer review skills;
- Performing corrections via revised versions;
- Publishing of all research findings and conditions.

Concerns with preprints

Sarabipour et al., 2019 highlight three major concerns that researchers often raise against preprints. However, evidence supporting these concerns is often virtually non-existent, or based on a mis-understanding of how preprints actually function.

1. *Preprints can lead to my research being scooped.* A common misconception. Preprints provide an authoritative timestamp of a discovery, in public, and usually with a DOI, or digital object identifier. Such is far more secure than, for example, presenting new results at a conference. Any 'scooping' will be akin to plagiarism and scientific misconduct, and can be dealt with as such.
2. *Preprinting prevents publication.* Another misconception. Virtually all major publishers/journals, including Elsevier, PLOS, Springer Nature, PNAS, and Wiley are friendly towards preprints. They allow them to be posted without compromising the publishability of authors' work in their journals. Therefore preprints are an excellent way to give researchers an advantage by increasing their publication record and citations.
3. *Preprints have low visibility.* How is publishing more work openly and faster decreasing the visibility of your work? Preprints, and discussions around them, are currently in an exponential growth phase (Tennant et al., 2018), and this is not likely to stop any time soon. All the evidence for now indicates that preprints are good for you, while increasing your online attention and citations.

Many publishers also allow preprints to be updated with the peer reviewed or accepted manuscripts too (postprints, remember them?), often even at the time of acceptance. Check journal policies either on their websites or SHERPA/RoMEO to see what you are permitted to share. A 2013 study by Björk et al found that the 'Green OA' coverage based on the sharing of these accepted manuscripts (postprints) was only around 12%, although with a lot of disciplinary variation. These were often shared in institutional repositories, but also on personal webpages of authors.

[INSERT INFOGRAPHIC FROM HERE https://figshare.com/collections/How_to_make_your_work_100_Open_Access_for_free_and_legally_multi-lingual_/3943972]

A recent analysis of preprints from bioRxiv by Abdill and Blekhman demonstrated that, as of 2017, around two-thirds of preprints shared there ultimately end up in peer reviewed journals. Furthermore, articles shared as preprints tended to end up in 'higher impact' journals and receive more downloads.

Examples of preprint servers/platforms

Here, platforms that end with "rXiv" of some sort are pronounced "archive" - the X in the name is actually the Greek letter chi in these cases. Possibly the most important thing you will learn in this whole module...

- **Cogprints** - An electronic archive for self-archive papers in any area of Psychology, Neuroscience, and Linguistics, and many areas of Computer Science, Philosophy, and Biology.
- **ESSOAr** - An Earth and Species Science open archive.
- **ChemRxiv** - A preprint server for Chemistry.
- **SSRN** - The Social Sciences Research Network. Note that SSRN is now owned by Elsevier.

- **Open Science Framework Preprints** - A collection of preprint service providers that use the OSF Open Source infrastructure to support their communities. This also aggregates preprints from other platforms into one search engine.
- **PrePubMed** - A platform that indexes preprints from arXiv q-bio, PeerJ Preprints, bioRxiv, F1000Research, preprints.org, The Winnower, Nature Precedings, and Wellcome Open Research.
- **preLights** - A service from the biological community that highlights specific preprints.

A really useful initiative in this space too is ASAPbio. This is a researcher-driven non-profit promoting transparency and innovation in life science communication.

Several language- or region-specific preprint servers have also recently emerged. These include Arabixiv (Arabic), Frenxiv (French), and INA-Rxiv (Indonesian), all hosted by the Open Science Framework. Add AfricArxiv here too.

Furthermore, a number of 'overlay' services now exist on top of preprints, including journals and other commenting/discussion platforms such as the Open Journal of Astrophysics which is an arXiv overlay journal. Similar to the popular StackOverflow, services such as MathOverflow and PhysicsOverflow provide discussion spaces where people can ask questions about preprints shared originally on the arXiv. Both include a community peer review system and postgraduate-level discussion forum. In the natural sciences, there is biOverlay. And in mathematics, Discrete Analysis. The movement for preprints has also begun now in Chemistry!

An exciting recent initiative in this space is Peer Community In. The "Peer Community in" (PCI) is a non-profit scientific organisation that aims to create specific communities of researchers reviewing and recommending, for free, unpublished preprints in their field (i.e. unpublished articles deposited on open online archives like arXiv.org and bioRxiv.org). So far, there are five specific communities that operate here:

- Peer Community in Evolutionary Biology
- Peer Community in Ecology
- Peer Community in Paleontology
- Peer Community in Animal Science
- Peer Community in Entomology

But how do I know which copy I can self-archive?

Thankfully, there is a service out there to make your life a little easier, called SHERPA/RoMEO. This is a database of journal self-archiving policies, and is based on a colour code:

- Green: Journals allow authors to share pre- and postprints
- Blue: Journals allow authors to share postprints
- Yellow: Journals allow authors to share preprints
- White: Journals to not formally allow authors to share their work

ASAPbio also has a handy little educational resource for open licensing for preprints.

Further reading