PREPRINTS: CONTINUING THE CONVERSATION

Preprint platforms have been common in physics and mathematics but the preprint landscape is changing rapidly with new platforms emerging across various disciplines. This raises opportunities for discussion across communities and for all those involved: preprint platforms, journals, authors, funders and institutions.

COPE has facilitated this discussion previously via an earlier forum discussion and with respect to peer review. To continue the conversation, COPE has developed a new 'Preprints' discussion document that sets out some of the benefits and challenges associated with the posting of preprints.

We have received valuable feedback since posting the discussion document and we'd very much welcome your thoughts and feedback as we discuss the key issues further:
- What are the risks if a preprint with a potential impact on public health is interpreted as established evidence?
- Could subject-specific differences influence the adoption of preprint servers, particularly in clinical fields?
- If preprints have DOIs, does that qualify them as prior publications?
- What standards exist for correcting preprints? Who is responsible? Could changes link to final published articles?
- Could researchers post a preprint that has been rejected from a journal?
- Do preprint servers address publication bias?
- Do preprint servers increase or decrease transparency?
- Who is responsible for checking if appropriate permissions have been received for the data and figures in a preprint?
  What if there is an authorship dispute on a preprint?

This was discussed at the start of the COPE Forum on Monday 30 April 2018.

COMMENTS FROM THE FORUM (Monday 30 April 2018) – NOTE, Comments do not imply formal COPE advice, or consensus.

- There needs to be a responsible approach to posting preprints, clearly stating what they are. The mainstream media need to be educated about preprints—what they are good for and what they are not.
- There is a concern of mis-use of data—that authors will post their data on preprint servers and others will use it, or mis-use, or misinterpret it. What might publishers do if they promote and allow the posting of material on preprint servers in terms of protecting authors and their data? However, mis-use of published data can also apply to published research publications and to some extent is how science works. An advantage of preprints is that authors can choose when they believe their research is ready to be publicly available—ie, when it's ready for further debate.
- The reality is that because there is a preprint mechanism in place, researchers will use it and so the important factor is to clearly label material as a preprint version, as distinct

from a 'peer reviewed' version. The focus should be on educating the media and the public on what peer review means, and how to distinguish between peer reviewed and non-peer reviewed material. Hence preprints should be clearly labelled with an explanation of what they are.

- Clear labelling is as an opportunity to demonstrate the benefit of peer review and explain more clearly what the process of peer review brings to journals—what checks have and have not been done.
- There is so much information available on the internet, and it is important that researchers can share information. Preprints can accelerate the sharing of information among researchers. The challenges are when preprints are misinterpreted as the final paper. Hence labelling and media education are very important. Publishers may need to consider how preprints are labelled as such within references.
- Preprint servers need to take a role in clarifying that what is posted on their site has not been peer reviewed. Perhaps peer reviewed papers should also be labelled (by publishers?) so that they can be distinguished from non-peer reviewed material, such as preprints. Should publishers be more explicit about the peer review process that the paper has undergone, what that process involves and what peer reviewers and editors are looking for?
- What are the implications of preprints for journals in terms of the objectivity of peer review, if reviewers are aware of the paper through a preprint (if peer review is double blind)?
- In terms of barriers to sharing data openly—the concerns of downstream use or fitness for use—this has been in discussion by the Research Data Alliance Legal Interoperability Group. From an institutional perspective, one consideration is that the General Counsel or Office of Tech Transfer requires disclaimers in relation to data accuracy for any release of data created at that institution. If the data are being produced by an institution that has an IP policy, this may be the first place to start, to ensure that institutions themselves are complying with the set of expectations around data accuracy and labelling.
- Everybody needs to be involved in these discussions—from preprint servers to publishers to institutions.
- Are preprints prior publication? If a preprint has a DOI, does that mean it can be considered as prior publication?
- Equating the assignment of a DOI with publication devalues the work of journals. Assigning a DOI so that there is a persistent link to the material should be encouraged in terms of scholarly record keeping but it should not indicate prior publication.
- Assignment of a DOI should not equate with consideration of material as a publication— examples include datasets/reviews, which are different from research article publications.
- Preprint DOIs can be linked to the article DOI of the final published paper (this was set up by CrossRef). Having a DOI means that preprints are citable. Should publishers have policies on citing preprints? Should journals include details of what has changed from the preprint to the published article?
- Assigning a DOI to a preprint does not mean prior publication. There is a terminology issue here. If something is published on the internet and is publicly available, then in the usual sense of the word it is considered published. However, preprints are preliminary versions of a paper and many journals have policies that do not preclude the publication

of the subsequent peer reviewed paper. Giving something a persistent identifier is a necessity but does not alter its publication status.

- arXiv does not assign DOIs to preprints. BioRxiv has recently added "version notes" to highlight changes: https://twitter.com/cshperspectives/status/984751893284089856
- Can we reverse the question and ask: in what way are preprints NOT publications? That might give us more clarity?
- Other terminology is type 1 versus type 2 "publication" https://pdfs.semanticscholar.org/9053/fce4ada0deea76705b97ac051dd581d91f82.pdf the difference between the two being peer review.
- Publication seems to carry different meanings—for example, publication for the purpose of copyright or national deposit versus making something publicly available. Theses suffer from this conflation also.
- Publication (capital P) is about curation and value adding; preprints are being made available (publication small P).
- Should preprints be considered in the same way as conference proceedings? Computer Science conferences are fully refereed on whole papers, not just abstracts, so they are not at all like preprints. Hence there are going to be cultural specific issues with preprints.
- The landscape of preprints is that there is no universal definition about who posts what when. The community at large needs to put some definitions and parameters around different types of preprints and then publishers can decide what preprint servers they may or may not accept for the subsequent publication of the peer review article. We need to define the different types of preprints.

ACTIONS: COPE will review the comments in relation to updating the discussion document in the future.

COMMENTS POSTED ON THE WEBSITE
Posted by Foppe van Mil, 24/4/2018
I can understand the critical position of the publishers well, unless you post the final paper (after peer review) as open access. Publishers current economic model includes 'selling' articles. With a paper already on a preprint platform, it is likely that other researchers use the preprint version. Additionally, such a preprint version has usually not been peer-reviewed, and may be faulty or erroneous from a scientific perspective. I would not consider a preprint document a solid contribution to the knowledge-base, and therefore would rather not see such platforms publicly accessible.

Posted by Gunther Eysenbach, 24/4/2018
You may want to add the JMIR Preprints platform to your list in the appendix: preprints.jmir.org. It covers all manuscripts submitted to JMIR journals and has an emphasis on technology in health care/medicine (direct submissions to the Preprint server is also possible).

Posted by Trevor Lane, 30/4/2018
Just some thoughts:
Ethics: Preprint platforms contain research information and are now part of the research/publishing cycle, so there need to be some governance and standards related to research

ethics, rather than just providing cloud server space for researchers to self-publish drafts. Some requirements would be universal such as disclosing conflicts of interest and funding, and all authors agree to posting; some would be discipline-specific such as human ethics approval, participant consent, and trial registration.

Public accessibility: Because preprints have not been peer reviewed but are part of the open research movement and are open access, there should be clear warnings to users who are not researchers (not just saying "not yet peer reviewed" as the public may not understand that; perhaps add "For research use only"). Or users could be restricted to people with an ORCiD account. However, research theses/dissertations are also not yet formally published and sometimes publicly available too. If the funder mandates open access or public domain status, then the preprint would have to adhere to that too.

Media use: Some concerns are that the mass media may not understand that preprints are for research use only and they may contain mistakes and exaggerations, or lack limitations and caveats; unlike journal articles, preprints lack press embargoes; unlike conference reports and presentations (ICMJE guidelines say don't give any more information to reporters), preprints may be the full and nearly final version; and the altmetrics trend may encourage researchers to widely publicize and share preprints.

DOIs: Having DOIs just gives a persistent link to any object, and are useful as a link in a citation in an article. To distinguish the preprint from a research article, a reference should say Preprint (as recommended by the NLM). But the in-text citation would need to say it too, such as "According to a preprint by X et al", rather than implying a published study. There would be opportunity to revise that if the preprint is published before publication of the paper, but if not, the preprint would have the DOI link to the published version. Linking would be a joint responsibility between publisher and preprint platform, with the latter also closing comments for the preprint version at that point. If there were any preprint comments that improve a paper between article acceptance and final publication, they could be used in a Correction.

Authors/Institutions need to know more about preprints as an additional avenue for research discussion, but also need to be aware of avoiding "research/science by press conference" and the impact of publicizing discoveries too early on patent applications.