PREPRINTS: WHAT ARE THE ISSUES?

Preprints and working papers have been posted and shared for many years. They report research results that have not undergone peer review, although in many cases the authors also submit to a journal (before, after or at the same time as making a preprint available). In the past 5 years, the number of preprint servers and preprints has expanded and new disciplines, notably biology and life sciences, have seen rapid growth in the number of preprints. Preprints appear in a number of contexts, including dedicated preprint servers (of which arXiv is by far the largest), social media and networking sites (e.g. Researchgate), and institutional repositories. Standards, pre-online checks and policies vary between platforms and there are currently no established standards, although ASAPbio (http://asapbio.org) will convene a meeting in July in which the establishment of standards will be among the issues discussed.

To date there have been few public discussions around the ethics of making unverified research available in this way and there are a number of issues that arise. Not all ethical issues around preprints have a link with journal articles and COPE may wish to consider 'mission creep' and whether all aspects of preprints ethics fall within its remit.

A few important issues for COPE to discuss are the following:

Citation – Can and should a preprint establish precedence for reporting research results? Should authors of journal articles be expected to include relevant preprints when reviewing the literature? Preprints are typically not included in regular indexing services, so are less visible than journal articles and more likely to be overlooked by authors.

Withdrawal/retraction and correction/updating preprints – Under what circumstances should a preprint be retracted or updated and whose responsibility is it to make such a decision, given that there is typically no expert editor associated with a preprint? If a journal version of a preprint is retracted, should the preprint also be removed or marked in some way? If a preprint is withdrawn should a reason for the removal be given? Since they have not been peer reviewed, preprints should be used with care by readers, however it is not clear whether there is a threshold at which leaving an inaccurate preprint online could be considered damaging and misleading.

Ethical approval – Could authors take advantage of preprint servers that have minimal checks to circumvent ethical requirements, e.g. for clinical trials? What recommendations can be made to those running preprint servers regarding ethical policies?

Journal editors – Many journals will consider papers that have been posted as preprints, although some do not have a clear policy. Are there any special considerations for papers that have already appeared as a preprint? For example:
- Consideration of online comments/reviews especially where they identify weaknesses of the paper;
- Verification that a previously posted version of a paper has not been peer reviewed;
- Whether to consider a submitted paper that reports results already included in a preprint with different authors.

Further reading
ASAPbio Draft statement 2: journal policies regarding pre-posting of articles:
http://asapbio.org/drafts/draft2.

Retraction/removal policies from several preprint servers:
- ArXiv https://arxiv.org/help/withdraw
- BiorXiv http://www.biorxiv.org/about/FAQ
- PeerJ Preprints https://peerj.com/about/preprints/policies-and-procedures/#retraction-policy
- Preprints.org https://www.preprints.org/instructions_for_authors#withdrawal

This was discussed at the start of the COPE Forum on 24 July 2017.

COMMENTS FROM THE FORUM (24 July 2017) – NOTE, Comments do not imply formal COPE advice, or consensus.

- It is imperative that disclosure of preprints by authors is made to journals. What use the journal makes of that information is up to the journal and its policies. COPE has always championed transparency and so it is extremely important to disclose any prior publication of work. It is then up to the editor to decide what to do or how to proceed, but the existence of the preprint should be disclosed. Should there be published standards for how this happens?

- Quality checks can vary and some preprint servers do not carry out quality checks. Some servers will do basic checks on quality after submission and these would not be expected to pick up image manipulation for example, or plagiarism. Should we expect preprint servers to do text comparisons (eg, iThenticate)? Could we develop such checks for journal/preprint co-submitted preprints, and so have a one stop shop? But preprint servers are not journals, and they do not carry out the types of checks that journals do, and perhaps they should not be expected to do these checks? Surely the value added by journals is to provide the necessary checks. It is great to have the research out there but there are no accepted standards at the moment. Preprint servers are not COPE members and so they have not signed up to the same standards and polices (eg, conflicts of interest statements, authorship criteria) as publishers and journals.

- There are no major issues with preprints if they are published. Preprints can be a very useful way to discuss research and make research better. Problems arise when the research remains a preprint and is not submitted to a journal, or if the preprint is not updated but people are still commenting on this, rather than the updated version. Also, what happens if the published version of the research is retracted? Should the preprint be retracted, or marked in some way? Should there always be a link between the preprint and the final published version?

- Can an article be submitted across several preprint servers? This does happen, where preprints have been submitted to servers that have been posted elsewhere. Some preprint servers issue DOI numbers, but others do not. This is a tricky issue and multiple submissions to preprint servers is not that common and is tolerated at the moment but it can mean that there are multiple preprints of the same text and so it is difficult to link to one preprint easily.

- From the journal's and the publisher's perspective, this is a complicated issue and a preprint policy would make sense.

- Readers need to understand the issues and the limitations of different forms of publication. The research needs to be cited as a preprint so that it is clear that the paper has not been peer reviewed and that other checks have not been made. Citation practices around preprints is one way that we can make it clear to readers of scholarly publications what the research is and make them aware of the fact that this is non-peer reviewed material.

- Some researchers think of preprints as scoop protection—protection against others posting the same work in a journal. If a preprint is out there with relevant work, should it be cited in the same way as a journal article, or is it too much of a burden on researchers to find work not in the major databases to make sure their literature review is complete?

- All authors should be in agreement with the preprint's content. Ultimately, once the paper is accepted for publication with more or less the same content as the preprint, the latter should be flagged to that effect (ie, that a peer reviewed version of the paper is now available).

- Education more broadly is key to this understanding. Preprints have been around for years but they were not widely disseminated (often they were departmental reports to stimulate discussion) and researchers understood the difference between these informal pieces of work and the published paper. It is in the interests of preprint servers to help in this education process.

- Authors need to consider the terms of licences and rights that are granted to preprint repositories and websites when they submit a paper and whether or not these might be in direct contradiction of the terms of the agreement with the publisher if their paper is then accepted for publication in the journal, particularly if the accepted article is not published under an open access creative commons licence. A lot of preprint servers put a license on the work, many give an option of a creative commons licence, while some have no licence. Should the final published work at the journal be considered the same or significantly different from the preprint version? In most cases, people consider it to be different, so there could be separate copyright agreements for each.

ACTIONS: COPE will consider producing a discussion document on this issue. Advice is needed on what policies journals should have to clarify the situation for authors and journals, and maybe even publishers, particularly in terms of whether or not preprints prejudice subsequent publication in a journal and if citation rules extend to preprints.

COMMENTS POSTED ON THE WEBSITE
Posted by Jordan Anaya, 18/7/2017
This discussion is long overdue as there have been dozens of life science preprints retracted for unknown reasons, via undisclosed procedures.

Here is an example of a preprint retraction at preprints.org:

https://www.preprints.org/manuscript/201703.0210/v1

The preprint simply states that it has been withdrawn. The abstract and PDF is no longer available (although you can get the abstract from PrePubMed).

BioRxiv has an even more questionable retraction procedure--the preprints are simply taken down, and the DOI just points to a "Page Not Found" page.

There needs to be a discussion about whether we should hold preprint servers to the same standards as journals when it comes to retraction notices.

Other important topics of discussion are the screening and retraction procedures at the various servers. A recent case highlights the urgent need for established standards:

A sole author posted 3 preprints on bioRxiv claiming a Cancer Cell paper was fundamentally flawed. It appears neither the author nor bioRxiv notified the Cancer Cell authors about the preprints. When the Cancer Cell authors found out about the preprints, and saw that they did not contain any meaningful scientific content, they requested them to be retracted. BioRxiv refused their request, so they resorted to simply commenting on the article. I blogged about this case a bit here: https://medium.com/@OmnesRes/crap-spotted-at-biorxiv-15eecd58be6f

This case raises several issues.

1. It is clear "non-scientific content" can get past bioRxiv's screening procedure. The question then becomes: what does bioRxiv do once it is discovered they mistakenly positively screened an article?

2. If the sole point of a preprint is to criticize another paper, should the author or bioRxiv be required to notify the authors of the criticized work to give them a chance to respond? Should they be notified prior to the posting of the preprint? Should bioRxiv perform a more careful screen of these articles since scientific reputations are at stake?

3. After my blog post the author of the preprints tried to respond via a comment on his own preprint, but his comment has been stuck in moderation. When bioRxiv does not approve a comment the commenter never learns why the comment was not approved, it simply stays in moderation forever. This is not the first time bioRxiv has moderated the comments on a controversial preprint.

Overall, I am growing increasingly concerned by the lack of transparency at bioRxiv. This includes why some preprints pass screening and others don't, why some preprints get taken down, and why some comments do not get posted.
Jordan Anaya, creator of PrePubMed

Posted by John Inglis, 21/7/2017
Responding to the issues raised here:
1. All content submitted to bioRxiv is screened through a two-step process. The first is done in-house and includes determination of relevance to the server and a plagiarism check. A second step is conducted by an international group of principal investigators with diverse expertise who are identified on the bioRxiv site. Anything posted on bioRxiv has been

approved as scientific content. The quality of that content is not relevant to the screening process. No "mistake" was made in the instance raised here.

2. The bioRxiv commenting function is open to anyone who wishes to make use of it. Notification of what is posted is impractical suggestion. (For example: what would be the threshold for notification? How could this be done with thousands of manuscripts arriving each month?) In the instance raised here, the authors of the criticized paper had as much time and space as they wished to rebut the criticism and they took advantage of it. The author of the critical paper also had that opportunity and used it to respond to the rebuttal. The exchange was open, unlimited, and is available to any reader who wishes to assess the merits of the arguments.

3. All comments on bioRxiv are moderated. This is a necessity to ensure that they are relevant to the topic of the manuscript involved and do not contain abuse or ad hominem comment. The particular comment you refer to is not "stuck in moderation". It was returned to the author with a request to resubmit it, focusing only on the scientific issues involved. It has not been resubmitted.

The bioRxiv FAQ contains information about the screening process and allowable material.

Comments are always posted unless they contain material that is offensive, personally abusive, or irrelevant to the subject of the manuscript.

Preprints on bioRxiv are removed only in situations in which legal issues require it. (For example, if an author has misrepresented some aspect of the submission, such as the agreement of co-authors or an academic affiliation). I have said publicly on several occasions, and to you personally, that there is a need for best practice in determining why a manuscript should be removed from a preprint server, how the removal should be indicated to readers, and whether or not any reason should be given. It is one of numerous instances in which cohesion on procedures among preprint servers in different disciplines would be valuable. There is the intention to have that discussion in a forthcoming meeting of server administrators.

Posted by Piotr Otręba, 19/7/2017
How journals which require original content, never published before, should actually treat pre-prints?

Posted by Anne Farmakidis, 21/7/2017
Our journal does not have a clear policy on preprints, but our manuscript submission process asks authors to disclose previous publications, etc. We recently had a situation that the author did not disclose that the submission had been published on a preprint platform. Although we think the author will be significantly transformed as part of the revision and editing process, we were troubled that the author did not disclose this previous posting of the submission. Curious how others have handled such issues.

Posted by Heather Tierney, 21/7/2017
Related to retracting/correcting preprints: is there ever a reason to post a Note/Expression of Concern on a preprint? What procedures/oversight would need to be used to post the concern?

Posted by Michael Wise, 23/7/2017
Sorry to be heretical, but are we perhaps over-thinking this? Preprint servers are, at best, informal publishers: places for lodging draft versions of articles for comment, replacing the old departmental reports series. Preprint servers are not COPE members, and, as things currently stand, would likely not be eligible for membership due to lack of appropriate policies. Authors therefore cite preprints at their peril, so the best remedy may be educational interventions about the role of preprints in the research publication landscape. In short, there is little COPE and its members can do about preprint servers, though it seems to me that the *rXivs aim to act ethically and are therefore to be preferred over the ethically more troublesome players in research publishing.

The issue for COPE members is whether to publish something that, in large part, as has already been made public via a preprint. This is not a new issue, and the resolution may turn on the degree of "value-adding" that comes with publication in a journal (not least the ethical quality control).

Posted by Charon Pierson, 23/7/2017
I agree with Michael here. Maybe one has to be old enough to remember the face-to-face discussions that used to occur among faculty and researchers prior to submission of articles (at least in some fields) where thoughtful commentary from one's peers helped shape the final paper. It seems that as everything else has moved online, so too has this function of pre-publication discussion, and in the process, there are unintended consequences. In the end, authors are responsible for citing appropriate literature and whether or not that includes pre-prints is likely discipline specific. I think editors and publishers need to develop policies appropriate to each journal.

Posted by Trevor Lane, 24/7/2017
Some issues that formal publications with formal peer review would need to consider are stating clearly in its information for authors if preprints are regarded as prior publication, if a declaration is needed at submission, and if preprints are allowed to be cited (with references labelling them as preprints or not peer reviewed). If a paper that was a preprint gets published, it should be compulsory for the authors to label the preprint as such and link to the final version, because there are often extensive revisions, corrections in data, and modified conclusions after peer review, such that the preprint is wrong. That would also stop people from wasting time informally reviewing the preprint. If a published paper gets retracted, the preprint should be marked/tagged as retracted too and linked to the journal retraction notice. Relevant policies are needed at least on the journal side in the form of author advice, especially as some journals are treating preprints as an option for green open access/archiving or self-archiving (yet, the preprint's content may be very different from that of the accepted version).