

Biomedical Journals and Preprint Services: Friends or Foes?

Moderators: Thomas Annesley^{1*} and Mitchell Scott²

Experts: Hilda Bastian,³ Vivian Fonseca,⁴ John P.A. Ioannidis,⁵ Michael A. Keller,⁶ and Jessica Polka⁷

As authors know, the current path for recognition and “scorekeeping” of contributions to biomedical science is publication in journals following the peer-review process. Many believe that the peer-review process is as old as scientific publishing itself. In reality, the peer-review process, as we know it today, is relatively recent. When the first scientific journal, the Royal Society’s *Philosophical Transactions*, began in 1665, the content was under the control of a single editor who solicited and selected papers he felt relevant and important. As scientific publications increased along with the number of journals, the review process remained under the control of single editors or at best a small team of editors until the early 19th century when some, but not all, journals would seek the advice from outside experts or referees about studies the editorial team was not familiar with. The peer-review process as we know it today did not really start until the middle of the 20th century. Indeed, *Nature* did not establish a formal peer-review process until 1967, and of Albert Einstein’s 301 publications there is evidence that only 1 underwent peer review. Interestingly, he told the editor of that journal that he would take his study elsewhere!

The advantages and flaws of the peer-review process have been debated and discussed in many formats, but this is the process for biomedical research publishing today. However, a movement to change the way biomedical research is reported appears to be burgeoning. This change, which is gaining publicity and support, is the immediate publication (“preprints”) of scientific results online without peer review on sites such as bioRxiv. While the preprint model of publishing online is new and foreign to the biomedical world, it has been the modus operandi for the physical sciences for over 20 years. Those in favor of preprint publication argue that sci-

entists are remiss at not publishing their results immediately as the ability to do so has been in place for over 2 decades (the Internet) and that delaying important biomedical findings is negligent to the public who fund the research.

The review process for scientific publications is an evolving one with the current peer-review process being the most recent iteration. Prepublishing may simply be the progression of this process. In February 2016, a group of well-esteemed biomedical researchers met at a workshop entitled Accelerating Science and Publication in Biology (ASAPbio) to discuss the potential role of preprints in facilitating the communication of biological research. Subsequently, in May 2016, the ASAPbio Funders’ Workshop was held at the NIH, at which scientists, existing preprint server managers, and funding agencies discussed the development of a preprint service for the life sciences. While the preprint movement is gaining momentum, there remain questions about preprints as a way of disseminating scientific research. These include the potential for “junk science” being posted on preprint sites, how and if preprints should be cited, how benchmark indices of scholarly achievement such as the *h*-index might be affected, the cataloguing of preprints (e.g., a “PrePubMed”), and how department chairs might “keep score” by using preprints for their research-engaged faculty.

Here we discuss the concept of preprints with a group of panelists representing various segments of the research and publishing community.

The original impetus for preprint servers was the long delay of the peer-review process in the physical sciences. In contrast, the time from submission to pub-

¹ University of Michigan Health Science Center, Ann Arbor, MI; ² Washington University School of Medicine, St. Louis, MO; ³ Lead editor for PubMed Commons, the commenting platform on PubMed, and PubMed Health, clinical effectiveness resources. Also academic editor at *PLOS Medicine*; ⁴ Editor-in-Chief, *Diabetes Care* (2006–2012), Tullis-Tulane Alumni Chair in Diabetes, and Chief, Section of Endocrinology, Tulane University, New Orleans, LA; ⁵ Professor of Medicine and of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, Professor of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA, and Associate Director, 8th International Congress on Peer Review and Publication; ⁶ Ida M. Green University Librarian,

Director, Academic Information Resources, Publisher, Stanford University Press, and Founder, HighWire Press; ⁷ Director, ASAPbio, and Visiting Scholar, Whitehead Institute, Cambridge, MA.

* Address correspondence to this author at: University Hospital, Room 2G332, 1500 East Medical Center Dr., Ann Arbor, MI 48109. Fax 734-763-4095; e-mail annesley@umich.edu.

Received October 13, 2016; accepted November 4, 2016.

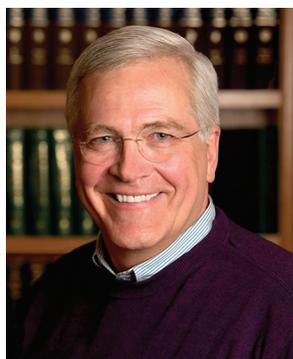
© 2016 American Association for Clinical Chemistry

lication for biomedical journals is generally much shorter. Including the publish-ahead-of-print (early posting) that most biomedical journals offer, what do preprints add that would not be accomplished by publication in a peer-reviewed journal within a few months of a preprint being posted?



John Ioannidis: The time from submission to publication for biomedical journals is indeed shorter than for journals in the physical sciences on average, but there is very large heterogeneity. Some biomedical fields and types of papers (e.g., methods) may have very long peer-review processes. Moreover, some

papers may have a hard time to get published despite representing well-done and informative work. This is particularly true of the most competitive papers, those that get submitted to the top journals. Since these top journals have acceptance rates of 2%–10% for uninvited noncommissioned submissions, many great papers may get repeatedly reviewed and rejected before finding a safe harbor. Other papers get accepted after several rounds of extensive revisions in the same journal. Thus having the preprint posted allows the wider scientific community to see/use the paper early on.



Michael Keller: I can think of several advantages to adding preprints to the publication cycles involving biomedical journals. First, exposure of a first-preparation manuscript to general view and potentially suggestions from members of the global community could bring unsuspected sugges-

tions and, ideally, even optimistically identify distant and unknown colleagues. Second, while the general statement that the length of time between submission and publication in biomedical journals is shorter than in other fields is true, there is great variation across journals. Third, making biomedical preprints widely accessible before and during peer review and the other important contributions of publishers serves to alert colleagues and other readers to one's findings in advance of publication, though certainly the final version of an article would be the one most people in the field would regard as definitive.



Jessica Polka: Any delay in making work available to the scientific community is a lost opportunity for collaboration, feedback, and cross-pollination, and this is especially true in fast-moving fields with public health relevance, such as Zika. Moreover, because many papers are submitted to multiple journals

sequentially to find a home, the time from submission to publication listed may be an underestimate of the total time a paper has spent under review.



Vivian Fonseca: Actually very little, if one considers published meeting abstracts as today's version of the preprint. Very few papers present important enough breakthroughs in medicine that will transform how we do things. Most of those that do usually come to light through abstracts and public announcements.

This is very different from the physical sciences.



Hilda Bastian: It's called "preprint," but it doesn't always precede an actual paper. Sometimes people want to make some work accessible, and get feedback on it, that may never be a paper in that form. It might be analogous to giving a conference presentation along the way. And sometimes, even a few

months could make a difference—especially when there is no guarantee that the first journal it is submitted to would accept it.

Are there changes or improvements to the preprint posting and archiving process that could help preprints become more accepted by the biomedical community?

Hilda Bastian: There are lots of unresolved questions about the pros and cons of preprints in most scientific communities, and perhaps especially in biomedicine. I'm

not sure all these issues can be resolved without becoming more like journals. The scale in biomedicine is vast: in quantity, breadth of subject matter, and types of studies and publications. Each extra level of safeguard against problematic papers, health claims, and conflicts of interest adds to editorial time and costs. This blurs distinctions between preprint repositories and journals.

John Ioannidis: ArXiv has a long-standing history and the lessons from its success can be safely transplanted in the equivalent efforts launched in the biomedical community. The main obstacle is not so much on the side of the preprint posting as on the side of the community being familiarized with the process and with all major journals accepting that preprint posting does not exclude a paper from being subsequently considered for publication in their pages.

Michael Keller: Discovery of relevant preprints is the key. Most commercial secondary publications (index and abstracting services) and aggregator sites (those commercial services assembling into one metadata domain information permitting readers to link or otherwise find articles and other genres) often do not include preprints.

Jessica Polka: I'm very excited for CrossRef's new preprint service, which will enable publishers to create a special type of DOI for preprints so that they can be better tracked through multiple versions.

More broadly, the preprint ecosystem in the life sciences is currently fragmented. There are many preprints in the q-bio section of arXiv, bioRxiv, and PeerJ Preprints, as well as other venues like figshare, institutional repositories, and preprints.org. While this diversity is great for giving authors a choice of the features and exposure they want for their work, we also believe there should be some common standards for preprints, and that they should be collected, archived, and made available for data mining in one place. Toward this end, ASAPbio is proposing the creation of a "central service" in response to a request from funding agencies.

Who holds the copyright for the final article: the preprint site or the scientific journal?

John Ioannidis: There are different ways to do this. Preferably, the author holds the copyright during the preprint posting, and copyright may then be transferred to a journal when the paper is published there—or one may follow whatever policy the specific journal has set.

Michael Keller: My advice to authors is to retain copyright for themselves and license publishers with negotiated parameters of engagement to permit publication.

Hilda Bastian: I am a strong believer in Creative Commons licensing. I don't think authors of scientific papers should give up copyright of their work. Many of us work under the aegis of a public agency where copyright isn't a question.

Jessica Polka: I don't know of any preprint server that requires that the authors transfer their copyright—hence, they are free to transfer it to a journal later.

Vivian Fonseca: This is a tough question. Could not journals or journal families function as preprint sites as well?

Are there any data available showing what percentage of preprints end up being published in peer-reviewed scientific journals?

Jessica Polka: A 2014 study published in *Journal of the Association for Information Science and Technology* found that among preprints posted to arXiv between 1995 and 2006, 73% could be matched to a published journal article in Web of Science. According to Paul Ginsparg during his keynote at the ASAPbio meeting at the Howard Hughes Medical Institute in February 2016, many of the remaining works are conference proceedings or theses. This suggests that most authors use both preprints and journals.

John Ioannidis: Data exist that compare arXiv and Web of Science-indexed journals, showing that 64% of arXiv preprints end up being published in Web of Science-indexed journals. There is substantial variability across disciplines though. Typically, the journal publication happens even in the same year as the preprint deposition or in the following year. For biomedical preprints, it is unknown whether the same pattern will be seen. Data from this early phase of biomedical preprints may be difficult to generalize yet.

Hilda Bastian: Most seem to be. A 2014 article in the *Journal of the Association for Information Science and Technology* by Vincent Larivière and colleagues found that 60% of arXiv papers were published in journals. The rate was increasing, and the time between preprint and paper was shrinking.

Vivian Fonseca: I am not aware of any. However, there are many abstracts in meeting proceedings that are never published or undergo peer review, so this will likely not be much different. I am usually very skeptical when such unpublished abstracts are cited or show up in review articles 3–4 years later. I would feel similarly about preprints that are never formally published.

How should preprints be listed on one's curriculum vitae? Should they be given the same weight as a peer-reviewed publication by department chairs or promotion review committees?

Vivian Fonseca: If the biomedical preprints don't end up as peer-reviewed publications within a year they are unlikely to be worth much. In my opinion, preprints should be given the same space and credit as meeting abstracts.

John Ioannidis: Papers should be assessed and judged on a case-by-case basis. Some preprints may achieve large citation impact and may clearly contain extremely important work. Experience from arXiv shows that there are preprints that have already received over 7000 citations according to INSPIRE, the High Energy Physics information system. Preprints on timely, important, and highly recognized work may even have more weight than the vast majority of peer-reviewed papers. Other preprints may be trivial, abandoned, refuted—but the same may also apply to peer-reviewed papers. A perusal of the bioRxiv preprint registry as of September 3, 2016, showed 5599 preprints. Any reader can judge the quality of this work but my preliminary impression is that much of it is of very good and excellent quality.

Michael Keller: Preprints might be used as placeholders in a curriculum vitae, but should not be given any weight until actually published in a peer-reviewed publication. Nor should the preprint get heavy consideration in any appointment, promotion, or tenure process.

Jessica Polka: Preprints can play an important role in evaluation because they showcase the applicant's most recent work. In situations that focus on recent performance, removing the delays associated with the vagaries of the peer- and editorial review process is critical—especially for junior researchers who may have few publications.

Of course, preprints are not peer-reviewed and lack a journal title, which conveys information about the stringency of peer review and the perceived importance of an article. However, they do, in principle, contain all the information required for a colleague in the authors' field to make a judgment about the quality of the work—provided a search or promotion committee invests the time to read the manuscript. While impractical at the early stages of a search process (for example, when there are 400 candidates under consideration for a job) reading manuscripts should not be out of the question for promotion committees. Indeed, others have called for this practice to combat other biases based on journal impact factor. Nevertheless, because curation and review by journals are an important part of our scientific evaluation

system, scientists who post preprints also need to submit those manuscripts to journals as soon as possible.

What should happen to a preprint once the manuscript undergoes peer review, revisions including re-analysis or addition of new data and the final article is published in a biomedical journal? Are mechanisms in place by preprint services to check for discrepancies between the preprint and the final published article?

John Ioannidis: Preprints can (and should) evolve as they improve with more work from the investigators and/or improvements suggested from peers. It is always nice to have a link to the respective paper eventually published in a journal. If there is continuity and cross-linking, having multiple early and intermediate versions should not be a problem. In the future, even papers published in peer-reviewed journals may commonly evolve/change after their publication, e.g., with postpublication review becoming more prominent.

Hilda Bastian: Ideally, if you post a preprint, it should be possible, at the very least, to link the published version to it. Reporting differences seems to me to be the authors' responsibility at this point. It would be great if they did it, and did it well, but realistically, that's not likely to happen consistently.

Michael Keller: In my opinion, the preprint should disappear because it is not a substitute for the final version.

Jessica Polka: Authors can update preprints after posting them, but many individual journal policies on this practice are currently vague. It would be possible to compare versions if the extracted full text of preprints and articles were both readily available. Unfortunately, PDFs remain the de facto standard for preprints in the life sciences.

Vivian Fonseca: I doubt there will ever be mechanisms to check for discrepancies—which is why the final article should trump the preprint once it is out.

What processes do preprint servers have (or need) to prevent fraudulent, fabricated, or just bad science from being posted? Similarly, what safeguards are in place to ensure that plagiarism or duplicate publication is not present?

John Ioannidis: Competent software can detect automatically in no time plagiarism or duplicate publication. More widespread use of such software by journals has led to a reduction in these phenomena. Still, some covert duplication cannot be fully excluded (e.g., the authors testing their paper for plagiarism with such software and then making amendments and edit so as to not be de-

ted). Similar processes can apply also to the preprint level. Fabricated science is very difficult to detect at the preprint level, or even in papers published in journals. There are some statistical forensic tools that can detect some fabricated papers or corpora of multiple fabricated papers. However, most fabricated work will not have any prominent features that can readily differentiate it from nonfabricated work. As for bad science, journals are full of it already and preprint servers will be just as well, unfortunately. Many efforts are underway to try to reduce research waste.

Michael Keller: The same questions need to be asked of publishers who basically publish articles on receipt of authors' publishing charges. Some sort of oversight is needed to assure that real science is submitted, that there is no plagiarism (this can be accomplished algorithmically), and that there is no duplication.

Hilda Bastian: I would add here the problems of undisclosed conflicts of interest, and one in biomedicine that is greater in magnitude than it is in many communities: of papers that trigger things like the autism-vaccine belief.

Jessica Polka: Preprint servers use a combination of automated and human screening to check for plagiarism and material that is not scientific in nature. However, aside from screening for content that is not a fair disclosure of scientific work or a violation of ethical standards, preprint servers do not perform any kind of peer review before posting.

Vivian Fonseca: We will have to rely on the same ethical principles we rely on for major society/conference abstracts, namely that the work is original.

Will preprint servers have an impact on "pay-to-publish" open-access journals?

John Ioannidis: Possibly, but the exact extent is unknown.

Hilda Bastian: I do not know.

Michael Keller: Possibly. The Cornell arXiv proponents seem to indicate, I recall, that the influence of that service on traditional publishing in the fields covered was mainly to improve the articles by early and widespread exposure to scientists in these fields.

Jessica Polka: I don't expect there will be a major impact on journals. Journals provide a "seal of quality" through organized peer review and editorial selection of varying stringency. While not everyone agrees on which journals are respectable or valuable, scientists will still be moti-

vated to seek out validation by peer review. This is essential for long-term career progression, and is not offered by preprint servers.

Vivian Fonseca: Probably not, unless journals start charging authors for preprints.

What's to prevent large publishing houses from commercializing the preprint process by creating appealing servers and charging for the preprint service? Could they require that authors submit manuscripts to one of their corresponding journals in lieu of the preprint posting fee?

John Ioannidis: The business wars that may ensue are difficult to predict. However, if 1 or more not-for-profit, free preprint servers are established and become popular in the biomedical sciences just as arXiv has in the physical sciences, this will make it difficult for commercial players to charge for preprint posting fees. They would have to offer some other major advantage in return and it is not obvious what this could be.

Michael Keller: A single preprint service per field, in this case biomedical disciplines, is ideal. Multiple preprint services will be something of a nightmare for scientists. It is also possible that large publishing houses would not accept an article that had been exposed in a preprint server.

Jessica Polka: Commercial publishers are interested in preprints, as evidenced by the sale of the social sciences archive, SSRN (Social Science Research Network), to Elsevier. The position of ASAPbio is that preprints are a public good. We are in conversations with funding agencies to ensure that they remain free to access and submit for at least the near term.

Vivian Fonseca: The latter is not a bad option if preprint servers were associated with particular "journal families" and a preprint fee is charged.

How are libraries going to deal with the idea that preprints will not be accessible through current institutional search venues, or possibly not indexed by subscriber-based services such as PubMed or the ISI (Institute for Scientific Information)?

John Ioannidis: Preprint posting aims primarily to achieve maximum visibility to the wide scientific community. It is hard to believe that these preprints will not be visible unless closed, for-fee preprint servers dominate the field. However, in this case, they will be more like for-fee journals without review rather than what arXiv has been in the physical sciences.

Michael Keller: Libraries will have to include preprint services as part of their discovery environments. This will take a bit of clever programming, but is certainly possible. Scientists and subject specialists in libraries will need to keep one another and their students/patrons informed on the availability of preprint services.

Jessica Polka: Preprints are indexed by Google Scholar, and several other tools (OSF Preprints, search.bioPreprint, and PrePubMed) have appeared to help scientists find life sciences preprints. The proposed central service that ASAPbio is advocating for, mentioned above, would be accessible by a search tool.

Vivian Fonseca: Some meeting abstracts can be searched—so these could likely be found in a similar manner.

Hilda Bastian: I think they will use services that will emerge for this—like the Open Science Framework aggregator.

Will the peer-review process need to change to remain relevant in the era of preprints?

Hilda Bastian: The peer-review process needs to keep adapting for a variety of reasons, especially the massive growth in research outputs globally and the wide range of possibilities that the Internet enables. I think peer review, like science generally, will become more collaborative, for example.

John Ioannidis: We need to improve peer review. Preprints are just one part of a very long chain. Improvements in peer review can happen at prepublication, publication, and postpublication level. By making scientific work visible to wide scientific communities before its formal publication, preprint posting may have a favorable impact on the skills and opportunities of the scientific community to contribute to peer review.

Michael Keller: Particularly in biomedical disciplines, peer review is essential. There are dangerous liabilities in simply accepting as verified science whatever someone might submit. There are ideological issues to understand as well. Some authors, hopefully very few, might regard having their preprint available from a known date as cre-

ating primacy in one or several aspects of presumed monetary advantage. In my opinion, we need always to emphasize, support, and participate in peer review.

Jessica Polka: The peer-review process in its current state is completely compatible with preprints. However, preprints help scientists receive additional feedback from the community, which may add to the effectiveness of journal-organized peer review. Moreover, preprints facilitate innovative peer-review models, such as overlay journals, which collect and curate content that is hosted in other archives (for example, the mathematics journal *Discrete Analysis*). The process of disclosing scientific work by posting a preprint does not erode the value provided by journal-organized peer review, which validates the work. For more on the distinction between disclosure and validation, see Vale and Hyman's 2016 article in eLIFE.

Vivian Fonseca: Peer review will likely evolve to have 2 steps. The first will be a quick decision by the preprint manager to preprint, followed secondly by a move on to full review—although the main conclusions are already in the public domain.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors' Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC; M.G. Scott, *Clinical Chemistry*, AACC; V. Fonseca, *Journal of Diabetes and its Complications*; M.A. Keller, HighWire Press; J. Polka, ASAPbio.

Consultant or Advisory Role: None declared.

Stock Ownership: None declared.

Honoraria: M.G. Scott, *Clinical Chemistry*, AACC.

Research Funding: None declared.

Expert Testimony: None declared.

Patents: None declared.

Previously published online at DOI: 10.1373/clinchem.2016.268227
