To Preprint or Not to Preprint

Forest W. Arnold[1]*, DO, MSc; Kimberley Buckner, BA[1]; Alex Glynn[1], MA; Jessica Petrey, MSLS[2]; Jose Bordon[3,4†], MD, PhD

[1]Division of Infectious Diseases, University of Louisville, Louisville, KY; [2]Kornhauser Health Sciences Library, University of Louisville, Louisville, KY; [3]Washington Health Institute, Washington, D.C.; [4]George Washington University Medical Center, Washington, D.C.

*f.arnold@louisville.edu; [†]jbordon@dc-whi.org

What is a Preprint?

A preprint is a version of a scholarly document that authors post on an open platform, such as a preprint server, an open-access researcher profile at their institution, or another freely accessible platform. Preprint posting typically occurs before submission or peer review for publication in a scholarly journal. Despite the implications of the term, preprints are not necessarily destined for publication; some may never be published in traditional journals and will exist on preprint servers indefinitely. If the manuscript is published, however, the term "preprint" refers to the version of the work that authors may share on platforms independent of the journal in which the manuscript was published according to policies of the journal in question. The posting of preprints is a continually growing trend dominated by just a few servers (**Figures 1 and 2**), but there is neither an understood preprint etiquette followed by authors, nor a universal preprint policy practiced by medical journals.

What is a preprint server?

A preprint server is a centralized online resource, network archive, or repository containing works or data associated with various scholarly papers that are not yet peer-reviewed or accepted by traditional academic journals. Such servers are predominantly free; 64 percent are owned by non-profit organizations.[1] Community-owned or scholar-led initiatives are the dominant model.

There are more than 60 preprint servers online.[2] A 2020 study identified 44 preprint servers with biomedical or medical scope.[1] Of these servers, 55 percent were multidisciplinary, while 45 percent accepted manuscripts in a single discipline only. Certain platforms also had restrictions concerning funding source, geographic region, and relevance to affiliated jo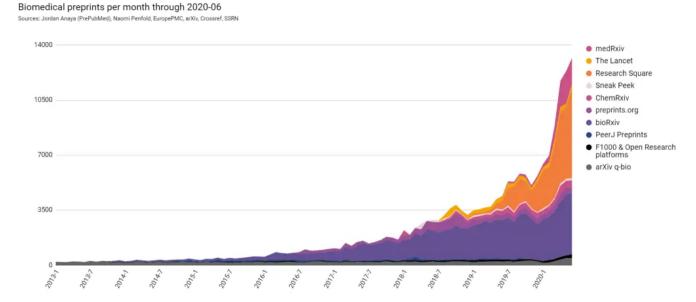urnals. The ASAPbio directory lists 55 preprint servers relevant to the life sciences, biomedical, and clinical research, indexed on 16 external sites.[3]

Why did preprints start in the first place?

Traditionally, several copies of a manuscript were snail-mailed to a journal. The journal would send these copies to reviewers, who would then mail them back to the journal. The journal would, in turn, eventually get the reviewer's comments back to the author. Typically, changes would have to be made and another version mailed back to the journal, who would consider it for publication. E-mail improved the speed of this process substantially—almost too well. It is not uncommon to be invited several times per day to review an article for publication. The proliferation of for-profit, online journals populating the market has only exacerbated the problem.

Two main factors prompted authors to preprint their manuscripts; both concern inordinate wait-times for publication, which for some journals can stretch from weeks to months or even years. First, authors were losing notoriety time for their work. Second, they were losing credit for a discovery or invention because the achievement was attributed, not to the person who did it first, but to the person but who published it first. Publishers have attempted to address these problems by advertising the submission-to-publication times of their journals. The designation "Epub ahead of print" was also created to allow authors' work to be disseminated prior to its official release date. But these efforts came too late to the negotiating table; the popularity of preprints and the number of preprint servers have already multiplied exponentially over recent years.

Biomedical preprints per month through 2020-06
Sources: Jordan Anaya (PrePubMed), Naomi Penfold, EuropePMC, arXiv, Crossref, SSRN

Figure 1. ASAPbio. Biology preprints over time. Available at: https://asapbio.org/preprint-info/biology-preprints-over-time. Accessed 5 April 2021.[4]
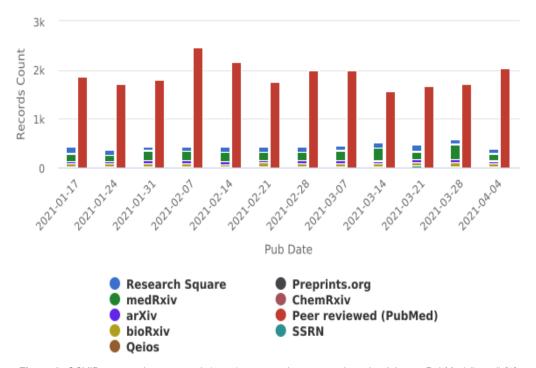
Figure 2. COVID-19 preprints per week (2021) compared to peer-reviewed articles on PubMed (in red).[5]

Is submitting a preprint to a journal for publication an ethical publishing practice?

The short answer is: it depends. There are no universal publication ethics, guidelines, licensing, or standard policies prohibiting authors from posting to preprint servers or servers from providing platforms for preprints. Thus, individual journals have their own policies regarding preprints. Most journals allow them, and some even screen preprint servers, looking for manuscripts to invite submission. Other journals prohibit the submission of any manuscript that has been posted on a preprint server. Journals that allow the submission of an article that has appeared as a preprint typically also require disclosure of the authors' preprint upon submission; not doing so constitutes a breach in the transparency that should exist between authors and journals. Authors and editors must be aware of the licensing implications of these practices, which vary by preprint platform.[6]

What are the benefits of a preprint for an author?

Early release facilitates early visibility, faster dissemination of an author's work, and due credit for a discovery or invention. Additionally, access to literature is more equitable, as most preprint servers do not have a fee structure; broader access in turn often leads to more citations. Since June 2020, PubMed Central has been piloting the indexing of preprints funded by NIH grants, a program explicitly intended to bring greater exposure to preprint manuscripts.[7] Some servers also enable feedback—conceptualized as 'crowdsourced' review [8] or 'informal peer review' [9]—and informal commenting on manuscripts; by incorporating this feedback, authors can submit a revised and improved version of their manuscript to a traditional journal with a better chance of acceptance for publication. A preprint effectively serves the same purpose as a poster placed at a national meeting, inviting early feedback on unpublished material.

A preprint is a complete manuscript, and though not peer-reviewed, it can be assigned a digital object identifier (DOI). Unlike URLs, which can be created at will by a private individual, DOIs must be registered with a DOI Registration Agency.[10] Consequently, the presence of a DOI can lend an article an appearance of legitimacy as a publication. DOIs also significantly increase the ease of citation; certain citation managers enable the user to automatically import every piece of information needed to cite a manuscript using only the DOI as input.

What are the negative aspects of a preprint for an author?

There are no industry-wide standards for the screening of preprints. The peer-review process of traditional academic publishing provides quality control and data validation for any published material. Given the lack of consistency across more than 20 categories of screening processes, informal feedback to authors is not an adequate substitute for traditional peer review.[3] Research shared online without undergoing rigorous quality control poses the risk of misinformation, misrepresentation, and misuse of data. Media reporting of low-quality research may ultimately be detrimental to patient care, public health, and the author's academic reputation.

Due to the ease of dissemination via preprint, authors may be tempted to rush into posting low-quality research on topical subjects, such as COVID-19. Additionally, any public comments made on a preprint may bias evaluations eventually performed by formal peer-reviewers. Some preprint servers do not allow commenting or provide limited functionality in commenting and feedback features, meaning that corresponding authors must rely on individuals contacting them personally by e-mail.

Authors should review their preferred journal's explicit or implicit policy of preprint acceptance before submission so that they do not invalidate their article's candidacy for publication. Many funders and universities also have policies that forbid the consideration of preprints in their assessment processes.

What are the implications for journals publishing preprinted manuscripts?

Journals ultimately want credit for publishing a manuscript, as citations of papers positively influence the journal's Impact Factor, an indicator of the average number of citations received per published article, which has historically been used to quantify the journal's prestige.[11, 12] Once a preprint is published in a journal, preprint servers should link the preprint DOI to the published journal citation and new DOI, but this does not always happen. It is the authors' responsibility to update servers with the publication status of their manuscripts. Unfortunately, there are not currently clear and enforceable ethical and retraction policies in place. Preprint servers have different policies for preservations and permanency once a document is posted; guidelines for removing a preprint from a server range from permitting an author to do so at will to restricting such removal to the server under extraneous circumstances.

Furthermore, since a manuscript may have the most

significant impact when it is first exposed to the media, a preprint submitted to a traditional journal for potential publication may already have been downloaded, discussed on multiple news outlets, and tweeted thousands of times. Since none of these communications would be associated with the journal potentially publishing the manuscript, they could be seen as lost opportunities for journal publicity, although some recent studies indicate that preprints may in fact expand a journal's reach.[13–15]

What is the bottom line?

The purpose of writing and publishing manuscripts is to distribute academic and scientific research, communicate the research effectively, and advance science and medicine. Manuscripts in the life sciences are intended to disseminate new knowledge to health care providers to facilitate improved clinical care and treatment. Preprinting non-peer-reviewed information runs the risk of favoring authors over patient safety. Prematurely released information may achieve positive outcomes in some cases, but it may also include potentially harmful misinformation that would have been corrected in the process of peer review. As Teixeira da Salva argues, "Preprints must always be understood as an incomplete work in progress."[16] Understanding that authors do not exist to serve publishers, authors may not feel obligated to share recognition with whatever journal eventually publishes their manuscript. Publishers do exist to serve authors, but preprints make this service complicated and difficult. Both authors and publishers should ultimately work to improve the health of society by disseminating accurate and relevant peer-reviewed information.

What is the policy of *ULJRI* regarding preprints?

We currently have a policy in place regarding duplicate submissions, which is similar to that of most other journals; it does not allow the submission of an article already under consideration at another journal. However, this policy does not cover the issue of preprints. As practice, expectations, and best practice guidelines evolve, we will continue to follow the ethical publishing practice of the Committee on Publication Ethics (COPE). We allow the submission of preprints, but we discourage authors from posting their articles as preprints, especially since we have no fees for authors or readers, and our turn-around from reviewers is relatively fast. For our COVID-19 special issue, we decreased the review period to just five days. We have expanded our policy regarding preprints to instruct authors to disclose that their submission was preprinted and provide the DOI and URL of all preprint versions. If we accept their manuscript for publication, we require that they notify all preprint servers carrying their article to update the citation. We also discourage the use of preprints as references because they have not been peer-reviewed. Any preprint cited must be labeled as such in the main text of the article and in the References section—guidance for referencing preprints is available on our Guidelines for Authors page. The full submission policy can also be viewed on our website.

Forest W. Arnold, DO, MSc   *Jose Bordon, MD, PhD*
Editors-in-Chief

Kimberley Buckner, BA   *Alex Glynn, MA*
Managing Editors

Jessica Petrey, MSLS
Review Editor

References

1. Kirkham JJ, Penfold NC, Murphy F, et al. Systematic examination of preprint platforms for use in the medical and biomedical sciences setting. BMJ Open **2020**; 10(12): e041849. doi: 10.1136/bmjopen-2020-041849. PMID: 33376175.

2. Johnson R, Chiarelli A. The Second Wave of Preprint Servers: How Can Publishers Keep Afloat? - The Scholarly Kitchen. Available at: https://scholarlykitchen.sspnet.org/2019/10/16/the-second-wave-of-preprint-servers-how-can-publishers-keep-afloat/. Accessed 16 March 2021.

3. Kirkham JJ, Penfold N, Murphy F, et al. A systematic examination of preprint platforms for use in the medical and biomedical sciences setting. bioRxiv [Preprint]. **2020**: 2020.04.27.063578. doi: 10.1101/2020.04.27.063578.

4. ASAPbio. Biology preprints over time. Available at: https://asapbio.org/preprint-info/biology-preprints-over-time. Accessed 5 April 2021.

5. NIH OPA *iSearch* COVID-19 Portfolio. Available at: https://icite.od.nih.gov/covid19/search/. Accessed 5 April 2021.

6. Council C. COPE Discussion document: Preprints: Committee on Publication Ethics, **2018** March 2018. doi: 10.24318/R4WByao2.

7. NIH Preprint Pilot. Available at: https://www.ncbi.nlm.nih.gov/pmc/about/nihpreprints/. Accessed 16 March 2021.

8. Ginsparg P. Preprint dj vu. EMBO J **2016**; 35(24): 2620–5. doi: 10.15252/embj.201695531. PMID: 27760783.

9. Pinfield S. Self-archiving publications. In: Gorman GE. International Yearbook of Library and Information Management 2004–2005: Scholarly Publishing in an Electronic Era. London: Facet Publishing, **2005**:118–45.

10. Digital Object Identifier System Handbook. Available at: https://www.doi.org/hb.html. Accessed 17 March 2021.

11. Abramo G, D'Angelo CA, Felici G. Predicting publication long-term impact through a combination of early citations and journal impact factor. J Informetr **2019**; 13(1): 32–49.doi: 10.1016/j.joi.2018.11.003.

12. Ranjan CK. Bibliometric indices of scientific journals: Time to overcome the obsession and think beyond the Impact Factor. Med J Armed Forces India **2017**; 73(2): 175–7. doi: 10.1016/j.mjafi.2017.03.008.PMID: 28924319 .

13. Fraser N, Momeni F, Mayr P, Peters I. The relationship between bioRxiv preprints, citations and altmetrics. Quant Sci Stud **2020**; 1(2): 618–38. doi: 10.1162/qss_a_00043.

14. Fu DY, Hughey JJ. Releasing a preprint is associated with more attention and citations for the peer-reviewed article. eLife **2019**; 8. doi: 10.7554/eLife.52646. PMID: 31808742.

15. Serghiou S, Ioannidis JPA. Altmetric Scores, citations, and publication of studies posted as preprints. JAMA **2018**; 319(4): 402–4. doi: 10.1001/jama.2017.21168. PMID: 29362788.

16. da Silva JAT. The preprint debate: What are the issues? Med J Armed Forces India **2018**; 74(2): 1620–4. doi: 10.1016/j.mjafi.2017.08.002. PMID: 29692484.