Educational Article

Adjusting the use of preprints to accommodate the ‘quality’ factor in response to COVID-19

Jaime A. Teixeira da Silva, PhD

Independent Researcher, Miki-cho, Kagawa-ken, Japan

Received 6 December 2020; revised 6 April 2021; accepted 8 April 2021; Available online 29 May 2021

Abstract

Preprints are typically crude precursors of peer-reviewed papers that are placed almost immediately, save for some superficial screening, on an open-access repository to allow the information to reach readers quickly, circumventing the long-drawn process typically associated with processing in peer-reviewed journals. For early-career researchers who might be enthusiastic about obtaining some recognition for their efforts, or wanting open and public input about their work, preprints are certainly a useful publication choice. However, if health-related data and information have not been carefully scrutinised, they may pose a risk and may even serve as a source of public health misinformation. Surging growth and competition among preprint servers, coupled with a massive volume of COVID-19-related preprints, mainly on bioRxiv and medRxiv, as well as select indexing now being tested on PubMed, suggests that preprints are being increasingly used in the biomedical sciences. Stronger and more robust ethical policies are needed to screen preprints before they are released to the public, and even if this implies a slight delay in publication, it may increase academics’ trust in this form of scientific information and communication. Clear and stringent ethical policies need to be urgently introduced by ethics groups such as COPE and the ICMJE, whose many member journals allow preprints to be posted before traditional peer review. Stringent ethical guidelines that treat misconduct equally in preprints and peer-reviewed papers will boost the integrity of academic publishing.

Keywords: bioRxiv; Ethics policies; Health policies; medRxiv; Peer review

© 2021 The Author.
Production and hosting by Elsevier Ltd on behalf of Taibah University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

E-mail: jaimetex@yahoo.com
Introduction: Battle for dominance of preprint market

Preprints are ‘a crude document representing information that has yet to be critically scrutinized by peers’ (p. 1026).\(^1\) The discussion about preprints continues to swirl around the issue of ‘quality’ control and the implications for citations of these non-peer-reviewed documents. A preprint is, by nature, a paper in a raw state of development and analysis. Following feedback from peers and the public before, or concurrent with, submission to a peer-reviewed journal, it can undergo alterations, general improvements, and enhancements, and can be gradually revised, evolving into new versions as it seeks a home in a preferably peer-reviewed journal. Alternatively, it can remain idle, as a preprint, with suspended fate, as a preprint, with suspended — and untested — ‘quality’. It is unclear how many preprints do not eventually make it into the peer-reviewed literature, or are not cited. These are subjects that are worthy of future analysis. However, some preliminary evidence already suggests that a solid percentage of published papers that were preceded by preprints garner higher altmetric scores than papers that were published without taking the preprint route.\(^7\) Fortified altmetrics, citations, and social media attention have been found for \textit{bioRxiv} and \textit{arXiv}\(^1\) preprints. Preprints are frequently touted as tools to promote the work of early-career researchers, many of whom see this form of publication as a rapid and easy way to showcase their developing work,\(^5\) or to enhance the speed and efficiency of scientific exchange.\(^6\) Apart from greater exposure and awareness of work via preprints, peer-reviewed papers that have passed through the preprint route ultimately appear to garner more citations.\(^9\)

This paper discusses the issue regarding the citation of preprints from a slightly adjusted philosophical perspective to that put forward by the author a few years ago.\(^7\) In that paper, and at that time, preprint platforms or servers such as preprint.org or \textit{bioRxiv} had just begun to take root, and the entire premise of the use of preprints within the publishing pipeline was still new and in a state of heated discussion and evolution.\(^10\) The notion of preprint ‘wars’ was proposed, suggesting that proactive competition by competing parties had begun for an expanding preprint market.\(^11\) Since preprints represent the first step in the gateway towards publication in a peer-reviewed journal, many of which are either for-profit subscription journals or open-access journals that charge an article-processing fee, the notion of preprint ‘wars’ — virtual competitive battles for an intellectual and financial reward — is neither unreal, nor far-fetched.

Not that long ago, the ‘acceptance’ of preprints among many mainstream journals and publishers was still not absolute; in other words, there was still a sector of the publishing community that actively rejected preprints, or that did not want to accept preprints for peer review, considering them to be ‘prior’ publications.\(^12\) Fast forward to 2019, about six months before the first cases of COVID-19 emerged in China, \textit{medRxiv} was launched by Cold Spring Harbor Laboratory (CSHL), Yale University, and BMJ as a preprint server for the medical sciences. The philosophy about the risks shared in 2017 has, in fact, not changed much, as certified by \textit{medRxiv} itself, which posts a notice on the top page of all preprints stating, ‘Caution: Preprints are preliminary reports of work that have not been certified by peer review. They should not be relied on to guide clinical practice or health-related behavior and should not be reported in news media as established information’. In other words, there are risks in using, citing, or relying on preprints as ‘documents of fact’. Some of those risks include the invalidation of claims or the lack of validation of claims, or the possibility of using or citing unscrutinised information, resembling ‘predatory’ publishing, whereby anything can be published or cited without scrutiny; the ease with which conflicts of interest can be hidden (or not); and the extended gaming of metrics.\(^13,14\) Despite the advances made in the preprint ‘market’ by \textit{medRxiv}, some members of the medical community continue to actively resist the acceptance of preprints related to clinical research for subsequent peer review.\(^15,16\) Their voices and concerns cannot, and should not, be ignored. Although preprints are gaining favour, a sector of academia remains resistant to this form of publication, the main issue being credibility, or the lack thereof.\(^17\)

Preprints in the COVID-19 era: Ethical challenges

Fast forward once again to 2021. With over 129 million cases and 2.81 million deaths worldwide due to the COVID-19 pandemic,\(^18\) CSHL’s \textit{bioRxiv} and \textit{medRxiv} have now become the ‘leading’ (in terms of volume) preprint servers in biology and medicine, with 14,682 preprints related to COVID-19 or SARS-CoV-2, the virus causing this disease, published to date (11,379 in \textit{medRxiv}, 3,303 in \textit{bioRxiv}).\(^19\) In some ways, almost ironically, COVID-19 ‘aided’ the projection of these two preprint servers to ‘success’ (i.e. increased use) and preprint market dominance. Other preprint servers or service providers, such as the suite of 26 preprint servers by the Center for Open Science (COS) hosting 2.297 million preprints,\(^20\) do not directly ‘compete’ thematically with \textit{bioRxiv}’s and \textit{medRxiv}’s preprint biology and medical science market prominence, which took just under eight years to be established (\textit{bioRxiv} launched in 2013), and their prominence was cemented in the COVID-19 era. An assessment of preprints in the first four months of 2020 indicated that 15% of abstracts in COVID-19-related papers underwent minor alterations by the time they had been published in peer-reviewed journals.\(^21\) A separate assessment of preprints in the first four months of 2020 found that the average time for a preprint to become a published paper was 63 days.\(^22\) Another study found that peer-reviewed COVID-19 papers took an average of 83.8 days between submission and publication, relative to 199.7 days for non-COVID-19 papers and 201.7 days for pre-COVID-19 papers.\(^23\)

There is a real and tangible risk of misinformation on human health, including exaggerated claims and hype,\(^24\) even more so now during the COVID-19 pandemic.\(^25\) That risk can emerge from preprints,\(^26\) peer-reviewed literature,\(^27\) or predatory publishing venues,\(^28\) both open-access or subscription, even more so given the deluge of COVID-19-related literature being published.\(^29\) One new and extremely serious risk in preprints is their silent (complete or partial) withdrawal or retraction from the published public record without any suitable explanation or transparent reason.\(^30\) If such papers were to be peer-reviewed literature, such
silent retractions/withdrawals would violate retraction policies by the Committee on Publication Ethics (COPE). However, preprint servers are not currently—as far as the author is aware—COPE members, and since COPE also apparently does not clearly adjudicate the ethics or provide ethical guidelines pertaining to the retraction of preprints, despite having a ‘position statement’ related to preprints, there is currently an ‘ethical vacuum’ pertaining to the ethics of preprint corrections and retractions that requires urgent debate, policy, and regulation, even more so now that lives are at stake during the COVID-19 pandemic.

However, many preprints are precursors of COPE member journals; in other words, many preprints end up being published in peer-reviewed COPE member journals. Therefore, the author is of the opinion that COPE has, through this association, the moral responsibility of adjudicating the ‘ethics’ pertaining to preprints, at least for its member journals and publishers, so as to maintain consistent ethical policies at the three main stages of the publication process: pre-publication (preprint), peer review, and post-publication peer review. The other important question that academics are surprisingly not asking is why none of the preprint servers (e.g., bioRxiv, medRxiv, COS’s 26 preprint services, preprint.org, etc.) have applied to be COPE members. One would envision that publishing integrity, as represented nowadays by COPE membership (journal or publisher), with due scrutiny and approval by COPE before the attribution of membership, would represent an important objective of such preprint servers, which should perhaps, as the leading preprint platforms, set the example of the future of preprint-related publishing ethics.

Other concerns about preprints

The risks and concerns about the use and citation of preprints have, in fact, not changed much in the past few years, even though both the number and volume of preprint servers, and preprints, have been increasing (Johansson et al., 2018), with almost 50% of bioRxiv preprints being published in Elsevier, Nature, PLOS, and Oxford University Press journals (Anderson, 2020). It is perhaps precisely because of their growth that concerns continue to exist. Given that they are citable items, and carry digital object identifiers (DOIs), involving a financial investment for DOI registration and the handling of processes related to preprint posting, online html text setting, indexing, and annual hosting costs, at least for COS preprint servers, supporters and owners of preprint servers understandably wish to promote them (even market them as brands) and their positive aspects in a ‘biased’ manner; they conveniently ignore the various risks and concerns raised in this paper and in others cited herein. The author considers the ‘promotion’ of preprints to be ‘biased’ because of an inherently skewed desire to ensure that invested efforts bear fruit. In this case, Naomi Penfold, the first author, was the ASAPbio Associate Director in 2018–2020, while Jessica Polka, the second author, is the ASAPbio Executive Director. In addition, PLOS, in which that paper was published, has a direct transfer agreement with bioRxiv and medRxiv, in the ‘Direct Transfer Program’ or B2J. Some of the most pertinent risks include promoting false or misleading information and a lack of clarity regarding policies between preprints and peer-reviewed journals. The risks of preprints to human health and publishing integrity may have been amplified by their indexing on PubMed in a National Institutes of Health (NIH) pilot programme managed by the National Library of Medicine (NLM), but exclusively for NIH-supported research, running for a minimum of 12 months, since June 2020. Allowing NIH-funded preprints to be indexed at PubMed, and its European affiliate Europe PMC, supported by high-ranking funding and health organisations (‘This COVID-19 preprints initiative is supported by a joint award from Wellcome, UK Medical Research Council (MRC), Swiss National Science Foundation (SNSF), and endorsed by the Chief Scientist of the World Health Organization (WHO)’), has led to an explosion of results when the term ‘preprint’ is searched for at PubMed (2,767 total: 240 documents prior to 2020, 1,924 documents in 2020, and 603 documents in 2021). Currently, since keyword-based papers are not allowed to be distinguished from the manuscript type (preprint), preprints have now in essence received a stamp of ‘validation’, qualifying them as citable equivalents of peer-reviewed papers at PubMed. Although this programme is experimental, a previous experimental post-publication peer review ‘pilot programme’ at PubMed (PubMed Commons) turned out to be a failure (Teixeira da Silva, 2018b). The graphs at Europe PMC also show a rapid increase in the preprint ‘market’ by Research Square, two prominent clients being Springer Nature and Cambridge University Press.

There are thus currently three risks of, or problems associated with, allowing preprints to be hosted at, and indexed by, PubMed: 1) giving them equivalent intellectual and citation status as peer-reviewed papers; 2) allocating a ‘validation’ stamp that allows them to be used for medical and public purposes, including for COVID-19, despite not being peer-reviewed; 3) providing an unfair advantage to US-based research that is funded by the NIH, even though PubMed indexes global research with researchers from around the world who also publish work in preprints, a policy that some may perceive as unfair.

Conclusion

The current debate surrounding preprints is as fervent as it had been a few years ago. In fact, given the prominence of COVID-19 in global health, society, and academia, preprints have now vaulted to the top rank of discussion topics in academic publishing. As biological, medical, and other academic literature continues to evolve, especially the corrective aspects of that literature, as preprints become more widely used, and as their volumes continue to rise, especially preprints related to COVID-19, the focus of the discussion is no longer ‘whether preprints should be cited’ but rather ‘how preprints should be cited’ to reduce the risks to the integrity of the literature and human health. Despite these concerns, to the author’s knowledge, there is currently no tangible proof that a preprint (as a document or as a source of information) is a risk, or threat, to human health, or that more harm to human health has been caused by
misinformation in a preprint than in a peer-reviewed journal. As much as peer-reviewed papers are supposed to be held to high standards, editors or managers of preprint servers should instil a basic, but rigorous, level of screening and quality control, at a minimum to establish open-science policies to ensure that basic ethical requirements are met.45 Bonnechère recommends that, to avoid possible harm, a preprint be removed after the final paper is published in a peer-reviewed journal.46 However, this is not advisable since a preprint represents a historical document that provides open and public evidence of the historical evolution of the paper, as well as a bibliometric record, and thus should never be removed. Ultimately, the integrity and credibility of preprints can be fortified only when they are treated as ‘ethical equals’ with peer-reviewed literature, for both COVID-19 literature47 and other fields of study.

Source of funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The author has no conflicts of interest to declare.

Ethical approval

The author confirms that this editorial had been prepared in accordance with COPE rules and regulations. Given the nature of the editorial, the IRB review was not required.

Authors’ contributions

The author contributed solely to the intellectual discussion underlying this paper, literature exploration, writing, reviews and editing, and accepts responsibility for the content and interpretation.

Disclaimer

Despite the author’s generally critical or skeptical views about preprints, he recognizes the importance and prominence of this form of publication, and even has four preprints (one each) at preprint.org, ResearchGate, and OSF’s PsyArxiv and SocArXiv.

References

1. Teixeira da Silva JA. Preprints should not be cited. Curr Sci 2017; 113(6): 1026–1027.
2. Serghiou S, Ioannidis JPA. Almetric scores, citations, and publication of studies posted as preprints. J Am Med Assoc 2018; 319(4): 402–404.
3. Fraser N, Momeni F, Mayr P, Peters I. The relationship between bioRxiv preprints, citations and altmetrics. Quant Sci Studies 2020; 1(2): 618–638.
4. Wang Z, Chen Y, Glänzel W. Preprints as accelerator of scholarly communication: an empirical analysis in mathematics. J Inform 2020; 14(4): 101097.
5. Wang Z, Glänzel W, Chen Y. The impact of preprints in library and information science: an analysis of citations, usage and social attention indicators. Scientometrics 2020; 125(3): 1403–1423.
6. Sarabiour S, Topf JM, Emmott E, Burgess SJ, Schwessinger B, Hensel Z. On the value of preprints: an early career researcher perspective. PLoS Biol 2019; 17(2):e3000151.
7. Poremski D, Falissard B, Fegert J, Witt A, Ordoñez AE, Martin A, et al. Moving from ‘personal communication’ to ‘available online at’: preprint servers enhance the timeliness of scientific exchange. Child Adolesc Psychiatr Ment Health 2019; 13: 42.
8. Hoy MB. 2020. Rise of the Rxivs: how preprint servers are changing the publishing process. Med Ref Serv Q 2020; 39(1): 84–89.
9. Fu DY, Hughey JJ. Releasing a preprint is associated with more attention and citations for the peer-reviewed article. eLife 2019; 8:e52646.
10. Teixeira da Silva JA. Teixeira da Silva JA. Preprint debate: what are the issues? Med J Armed Forces India 2018; 74(2): 162–164.
11. Teixeira da Silva JA. The preprint wars. AME Med J 2017; 2: 74.
12. Teixeira da Silva JA, Dobrânszki J. Preprint policies among 14 academic publishers. J Acad Libr 2019; 45(2): 162–170.
13. Teixeira da Silva JA. Preprints: ethical hazard or academic liberation? KOME 2017; 5(2): 73–80.
14. Teixeira da Silva JA. Intellectual phishing, hidden conflicts of interest and hidden data: new risks of preprints. J Adv Res Educ 2017; 4(3): 136–146.
15. Leopold SS, Haddad FS, Sandell LJ, Swiontkowski M. Editorial: clinical orthopaedics and related research, the bone & joint journal, the journal of orthopaedic research, and the journal of bone and joint surgery will not accept clinical research manuscripts previously posted to preprint servers. Clin Orthop Relat Res 2019; 477(1): 1–4.
16. Kharasch ED, Avram MJ, Clark JD, Davidson AJ, Houle TT, Levy JH, et al. Peer review matters: research quality and the public trust. Anesthesiology 2021; 134(1): 1–6.
17. Soderberg CK, Errington TM, Nosck BA. Credibility of preprints: an interdisciplinary survey of researchers. J Acad Libr 2019; 45(2): 162.
18. John Hopkins University. https://coronavirus.jhu.edu/map.html (last accessed: April 1, 2021).
19. medRxiv. https://connect.medrxiv.org/relation/content/181. (last accessed: April 1, 2021).
20. Center for Open Science. https://osf.io/preprints/ (last accessed: April 1, 2021).
21. Polka JK, Dey G, Pálfy G, Nanni F, Brierley L, Fraser N, et al. Preprints in motion: tracking changes between posting and journal publication. bioRxiv 2021. https://doi.org/10.1101/2021.02.20.432090 (preprint, not peer reviewed).
22. Seryugina YV, Dicks AJ. Publication practices during the COVID-19 pandemic: biomedical preprints and peer-reviewed literature. bioRxiv 2021. https://doi.org/10.1101/2021.01.21.247563 (preprint, not peer reviewed).
23. Horbach S. No time for that now! Qualitative changes in manuscript peer review during the Covid-19 pandemic. Res Eval 2021:rvaa037.
24. Ioannidis J. Coronavirus disease 2019: the harms of exaggerated information and non-evidence-based measures. Eur J Clin Invest 2020; 50(4):e13222.
25. Vlasschaert C, Trap JM, Hiremath S. Proliferation of papers and preprints during the coronavirus disease 2019 pandemic: progress or problems with peer review? Adv Chron Kidney Dis 2020; 27(5): 418–426.
26. Teixeira da Silva JA. Letter to the editor: editorial: clinical orthopaedics and related research, the bone & joint journal, the journal of orthopaedic research, and the journal of bone and joint surgery will not accept clinical research manuscripts previously posted to preprint servers. Clin Orthop Relat Res 2020; 478(9): 2186–2187.

27. Chirico F, Teixeira da Silva JA, Magnavita N. “Questionable” peer review in the publishing pandemic during the time of Covid-19: implications for policy makers and stakeholders. Croat Med J 2020; 61(3): 300–301.

28. Teixeira da Silva JA. An alert to COVID-19 literature in predatory publishing venues. J Acad Libr 2020; 46(5):102187.

29. Colavizza G, Costas R, Traag VA, van Eck NJ, van Leeuwen T, Waltman L. A scientometric overview of CORD-19. PLoS One 2021; 16(1):e0244839.

30. Teixeira da Silva JA. Silently withdrawn or retracted preprints related to Covid-19 are a scholarly threat and a potential public health risk: theoretical arguments and suggested recommendations. Online Inf Rev 2021. https://doi.org/10.1108/OIR-08-2020-0371 (in press). https://www.emerald.com/insight/content/doi/10.1108/OIR-08-2020-0371/full/html.

31. COPE (Committee on Publication Ethics). COPE Discussion document: preprints (version 1); 2018. https://doi.org/10.24318/R4WByao2 (last accessed: April 1, 2021).

32. Johansson MA, Reich NG, Meyers LA, Lipsitch M. Preprints: an underutilized mechanism to accelerate outbreak science. PLoS Med 2018; 15(4):e1002549.

33. Anderson KR. bioRxiv: trends and analysis of five years of preprints. Learn Publ 2020; 33(2): 104–109.

34. Penfold NC, Polka JK. Technical and social issues influencing the adoption of preprints in the life sciences. PLoS Genet 2020; 16(4):e1008565.

35. ASAPbio. https://asapbio.org/about-us (last accessed: April 1, 2021).

36. Public Library of Science (PLoS). https://plos.org/open-science/preprints/ (last accessed: April 1, 2021).

37. Sheldon T. Preprints could promote confusion and distortion. Nature 2018; 559(7715): 445.

38. Klebel T, Reichmann S, Polka J, McDowell G, Penfold N, Hindle S, et al. Peer review and preprint policies are unclear at most major journals. PLoS One 2020; 15(10):e0239518.

39. National Institutes of Health (NIH), National Library of Medicine (NLM). https://www.ncbi.nlm.nih.gov/pmc/about/nihpreprints/ (last accessed: April 1, 2021).

40. EuropePMC.org. https://europepmc.org/Preprints (last accessed: April 1, 2021).

41. PubMed. https://pubmed.ncbi.nlm.nih.gov/?term=preprint (last accessed: April 1, 2021).

42. Teixeira da Silva JA. PubMed Commons closure: a step back in post-publication peer review. AME Med J 2018; 3: 30.

43. Research Square. https://www.researchsquare.com/ (last accessed: April 1, 2021).

44. Teixeira da Silva JA. Evolution in the correction of the literature: preprints, manuscript versioning, error amendment, and retract and replace; 2020. https://doi.org/10.20944/preprints201708.0029.v2. Preprints.org (preprint, not peer reviewed).

45. Nieto I, Navas JF, Vázquez C. The quality of research on mental health related to the COVID-19 pandemic: a note of caution after a systematic review. Brain Behavior Immun - Health 2020; 7: 100123.

46. Bonnechère B. Preprints in medicine: useful or harmful? Front Med 2020; 7: 579100.

47. Teixeira da Silva JA, Bornemann-Cimenti H, Tsigaris P. Optimizing peer review to minimize the risk of retracting COVID-19-related literature. Med Health Care Philos 2021; 24(1): 21–26.