

The challenge of preprints for public health

O desafio dos *preprints* para a saúde pública

El desafío de los *preprints* para la salud pública

Kenneth Rochel de Camargo Jr. ¹
Cláudia Medina Coeli ²

doi: 10.1590/0102-311XEN168222

Preprints are “a form of a scholarly article which is not peer-reviewed yet but made available either as paper format or electronic copy” ¹. After an early attempt by the U.S. National Institutes of Health in the early 1960s, this format really took hold in the early 1990s, first as an email server at Los Alamos National Laboratory, which later became a web service known as *arXiv* ¹. In the following years, the number of both preprint servers and total preprints submitted to web services increased considerably, however, preprints are still a small fraction (6.4%) of the total output of scientific publication ¹.

Despite disagreements over whether this form of publication is actually beneficial or not, its advantages and problems present a high degree of convergence among advocates and detractors. On the one hand, preprint is beneficial because it is a quicker way to disseminate scientific content with open access to everyone; on the other hand, the lack of adequate vetting, especially for peer reviews, increases the risk of disseminating bad science and can lead to several problems ². The dissent lies in considering to what extent possible risks overcome possible benefits (or vice versa).

The argument about this rapid dissemination has strong supporting evidence. A study on preprint publication showed that preprint are published on average 14 months earlier than peer-reviewed journal articles ¹. This is expected considering that the time-intensive process of peer reviews and revising manuscripts is totally bypassed. However, in this strength lies its very fragility: how to assure that this shorter process will not compromise the quality of the publication?

ASAPbio (Accelerating Science and Publication in Biology) ³ is a group of biology researchers that promotes preprint publication and has produced a number of studies that attempt to allay concerns about its quality, claiming, for example, that published articles previously submitted to a preprint server did not show relevant changes for its publication ⁴. Authors from this group have argued that the current approaches to evaluate research and researchers hold back a more widespread adoption of the preprint methodology ⁵, which would explain its relatively small participation on the general panorama of scientific publication.

Despite claims to the contrary, however, there are examples of poor studies published as preprints, which caused undesirable consequences in public health. Two methodologically flawed studies about a protective effect of tobacco smoking against COVID-19 (one of which has an author with known connections with the tobacco industry), for example, increased the commercialization of tobacco products in France and Iran ⁶ and a virology study that erroneously stated that the SARS-COV-2 virus had “HIV insertions” fueled conspiracy theories about the former virus being a bioweapon,

¹ Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil.

² Instituto de Estudos em Saúde Coletiva, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

Correspondence

K. R. Camargo Jr.
Departamento de Planejamento e Administração em Saúde, Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro.
Rua São Francisco Xavier 524, Bloco D, 7º andar, Rio de Janeiro, RJ 20559-900, Brasil.
kenneth@uerj.br



which lingered on even after the preprint was removed from the server due to its egregious errors⁷. Studies have found that much of the public discussion and even policy was indeed driven by what was published in preprints rather than in scientific journals^{7,8,9,10}, thus, quality issues are a major cause of concern.

On the other hand, similar errors have been observed within traditional publishing; the publication of a poor quality paper with undisclosed conflicts of interest in one of the most prestigious medical journals, *The Lancet*, which became the trigger for the contemporary wave of anti-vaccine activism, is a major, and regretful, example. Understanding to what extent this problem is likely to occur with or without gatekeeping mechanisms is necessary.

Preprint advocates countered that the effect of poor science disseminated via preprints would be lessened by media reporting that explicitly indicated that those studies did not undergo any peer review and, thus, required more criticism and reserve before being considered essential sources for a public debate. It was probably the case of South African media⁸, but in Brazil, a study found that less than 40% of preprint-based reports on mass media clearly showed their provisional character¹¹.

A hypothesized advantage for preprints – namely open to wider discussion, as opposed to the limited criticism offered by editors and peer reviewers – does not seem to exist in real life. Most preprints do not elicit any comment, even considering a sharp increase in COVID-19-related articles (17.5% and 12.3% on *bioRxiv* and *medRxiv*, respectively, versus 3.2% and 1.4% for non-COVID-19-related material on both servers). Moreover, to date, comments were limited to one per article¹² and very few preprints showed any change since they were submitted to a server. Very few existed in more than two versions⁹. Much of the aforementioned problems resulted from scientific publishing, especially on biomedical sciences, which was overloaded with the COVID-19 pandemic^{13,14}.

The analysis of published articles about COVID-19, although they were proportionately small in comparison with the overall number of published articles, was considerably higher (0.097) than that of comparable publications on other world health emergency viral epidemics (from 0.023 to 0.024). This difference is even more concerning when considering that the time window of publications on other cases, such as Ebola or SARS, was considerably larger¹⁵. Similar rates of retractions were observed in preprint servers and journals, but for the former most retractions were related to ethical issues, which might have been noticed by editors and reviewers¹⁶. Still in the context of the pandemic, the proportion of preprints that were published in journals was far below than what was observed by Xie et al.¹ in a general context. The publication rate was only 5.7%, even allowing for a somewhat extended period of observation¹⁷.

The quality issues of preprints have been acknowledged even by its advocates and the major servers have adopted several measures to give some assurance to their users. Both *arXiv* and *bioRxiv* have advisory boards and rules for submission: the latter, for instance, checks for plagiarism and materials with offensive and/or non-scientific content that might pose health risks^{1,18} and added a warning about the preliminary nature of preprints, especially those about the pandemic¹⁹. Other repositories have enhanced their screening on articles related to the pandemic and *bioRxiv* no longer accept articles about COVID-19 treatments based only on computational work¹. Efforts to curate preprint materials have been proving to be effective regarding later acceptance for publication. Preprints included in the Centers for Disease Control and Prevention (CDC) COVID-19 Science Update had a much higher publication rate – almost two thirds if sufficient time allowed²⁰. However, does not adopting such measures bring preprint servers closer to the traditional publishing model? According to Maslove¹⁸ (p. 444), “to some extent, these measures are euphemisms for editorial boards and peer-review processes, raising the question of whether preprint servers differ from traditional journals in kind, or simply by degree”.

Accelerating the production of knowledge and expanding the scope of criticism over it are worthy goals, however, such goals must be pondered with possible risks of prematurely disseminating uncertain or even wrong information. As some sociologists of scientific knowledge accurately suggested, the timing of politics and the timing of science are to great extent different^{21,22}. The pressure to produce results in the face of a major public health challenge – such as the COVID-19 pandemic – is understandable, but allowing such pressure to compromise the very research outcomes society depends on to make proper decisions is counterproductive, to say the least. The implications of bringing out (relatively) unvetted materials are quite different for fields such as high-energy physics, in which the first successful preprint server arose, and those that directly affect the health of popula-

tions. During the last years, immense efforts to implement controls to avoid the publication of flawed clinical trials, for example, have emerged. Is it worth to sacrifice the necessary time to rigorously evaluate those in order to gain a few months in terms of publication?

Finally, cooperation between different research groups is the norm in contemporary science. Over the last years, networks of researchers tackling specific issues have been established worldwide. Information flows within such networks already exist regardless of the preprints, thus, the extent that those would add to such interactions is debatable.

Despite the criticisms, preprints are a reality in contemporary scientific publishing and can play a positive role in the production and dissemination of knowledge, as long as adequate care is taken to prevent bad or dangerous works to negatively affect the general public. The possibility of more widespread criticism of articles is a worthy goal and should be incentivized. This issue is linked to a wider discussion about peer review in scientific publishing, which requires its own debate.

Contributors

K. R. Camargo Jr. contributed to the study conception, writing, and review. C. M. Coeli contributed to the study conception and review. Both authors approved the final version to be published.

Additional informations

ORCID: Kenneth Rochel de Camargo Jr. (0000-0003-3606-5853); Cláudia Medina Coeli (0000-0003-1757-3940).

Acknowledgments

The authors thank the Brazilian National Research Council (CNPq; C.M.C.: 303295/2019-8 and K.R.C.Jr.: 306228/2021-1), the Carlos Chagas Filho Rio de Janeiro State Research Foundation (FAPERJ; C.M.C.: E-26/200.003/2019 and K.R.C.Jr.: E-26/202.893/2018), and the Rio de Janeiro State University (UERJ; Programa Prociência).

References

1. Xie B, Shen Z, Wang K. Is preprint the future of science? A thirty year journey of online preprint services. *arXiv* 2021; 17 feb. <https://arxiv.org/abs/2102.09066>.
2. Berg JM, Bhalla N, Bourne PE, Chalfie M, Drubin DG, Fraser JS, et al. Preprints for the life sciences. *Science* 2016; 352:899-901.
3. Accelerating Science and Publication in Biology. ASAPbio is a scientist-driven non-profit promoting transparency and innovation in life science communication. <https://asapbio.org/> (accessed on 06/Sep/2022).
4. Brierley L, Nanni F, Polka J, Dey G, Pálffy M, Fraser N, et al. Preprints in motion: tracking changes between preprint posting and journal publication during a pandemic. *bioRxiv* 2021; 14 oct. <https://www.biorxiv.org/content/10.1101/2021.02.20.432090v3>.
5. Penfold NC, Polka JK. Technical and social issues influencing the adoption of preprints in the life sciences. *PLoS Genet* 2020; 16:e1008565.
6. van Schalkwyk MCI, Hird TR, Maani N, Pettecrew M, Gilmore AB. The perils of preprints. *BMJ* 2020; 370:m3111.
7. Majumder MS, Mandl KD. Early in the epidemic: impact of preprints on global discourse about COVID-19 transmissibility. *Lancet Glob Health* 2020; 8:e627-30.
8. van Schalkwyk F, Dudek J. Reporting preprints in the media during the COVID-19 pandemic. *Public Underst Sci* 2022; 31:608-16.

9. Fraser N, Brierley L, Dey G, Polka JK, Pálffy M, Nanni F, et al. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biol* 2021; 19:e3000959.
10. Brierley L. Lessons from the influx of preprints during the early COVID-19 pandemic. *Lancet Planet Health* 2021; 5:e115-7.
11. Oliveira T, Araujo RF, Cerqueira RC, Pedri P. Politização de controvérsias científicas pela mídia brasileira em tempos de pandemia: a circulação de *preprints* sobre Covid-19 e seus reflexos. *Revista Brasileira de História da Mídia* 2021; 10:30-52.
12. Kodvanj I, Homolak J, Virag D, Trkulja V. Publishing of COVID-19 preprints in peer-reviewed journals, preprinting trends, public discussion and quality issues. *Scientometrics* 2022; 127:1339-52.
13. Bauchner H, Fontanarosa PB, Golub RM. Editorial evaluation and peer review during a pandemic: how journals maintain standards. *JAMA* 2020; 324:453-4.
14. Gianola S, Jesus TS, Barger S, Castellini G. Characteristics of academic publications, preprints, and registered clinical trials on the COVID-19 pandemic. *PLoS One* 2020; 15:e0240123.
15. Tentolouris A, Ntanasisis-Stathopoulos I, Vlachakis PK, Tsilimigras DI, Gavriatopoulou M, Dimopoulos MA. COVID-19: time to flatten the infodemic curve. *Clin Exp Med* 2021; 21:161-5.
16. Boschiero MN, Carvalho TA, Marson FAL. Retraction in the era of COVID-19 and its influence on evidence-based medicine: is science in jeopardy? *Pulmonology* 2021; 27:97.
17. Añazco D, Nicolalde B, Espinosa I, Camacho J, Mushtaq M, Gimenez J, et al. Publication rate and citation counts for preprints released during the COVID-19 pandemic: the good, the bad and the ugly. *PeerJ* 2021; 9:e10927.
18. Maslove DM. Medical preprints – a debate worth having. *JAMA* 2018; 319:443-4.
19. Heimstädt M. Between fast science and fake news: preprint servers are political. *LSE Impact Blog* 2020; 3 apr. <https://blogs.lse.ac.uk/impactofsocialsciences/2020/04/03/between-fast-science-and-fake-news-preprint-servers-are-political/>.
20. Otridge J, Ogden CL, Bernstein KT, Knuth M, Fishman J, Brooks JT. Publication and impact of preprints included in the first 100 editions of the CDC COVID-19 science update: content analysis. *JMIR Public Health Surveill* 2022; 8:e35276.
21. Eyal G. *The crisis of expertise*. New York: John Wiley & Sons; 2019.
22. Collins H, Evans R. *Why democracies need science*. New York: John Wiley & Sons; 2017.

Submitted on 06/Sep/2022
Approved on 08/Sep/2022