



Current concerns on journal article with preprint: *Korean Journal of Internal Medicine* perspectives

Hyoung-Chul Oh¹ and Hyun Kang²

¹Division of Gastroenterology, Chung-Ang University College of Medicine, Seoul; ²Department of Anesthesiology and Pain Medicine, Chung-Ang University College of Medicine, Seoul, Korea

Preprints are preliminary research reports that have not yet been peer-reviewed. They have been widely adopted to promote the timely dissemination of research across many scientific fields. In August 1991, Paul Ginsparg launched an electronic bulletin board intended to serve a few hundred colleagues working in a subfield of theoretical high-energy physics, thus launching arXiv, the first and largest preprint platform. Additional preprint servers have since been implemented in different academic fields, such as BioRxiv (2013, Biology; www.biorxiv.org) and medRxiv (2019, Health Science; www.medrxiv.org). While preprint availability has made valuable research resources accessible to the general public, thus bridging the gap between academic and non-academic audiences, it has also facilitated the spread of unsupported conclusions through various media channels. Issues surrounding the preprint policies of a journal must be addressed, ultimately, by editors and include the acceptance of preprint manuscripts, allowing the citation of preprints, maintaining a double-blind peer review process, changes to the preprint's content and authors' list, scoop priorities, commenting on preprints, and preventing the influence of social media. Editors must be able to deal with these issues adequately, to maintain the scientific integrity of their journal. In this review, the history, current status, and strengths and weaknesses of preprints as well as ongoing concerns regarding journal articles with preprints are discussed. An optimal approach to preprints is suggested for editorial board members, authors, and researchers.

Keywords: Preprint; medRxiv; Peer review; Research report

INTRODUCTION

Preprints are preliminary research reports that have not yet been peer-reviewed. They have been widely adopted to enhance the timely dissemination of research across many scientific fields [1,2]. However, because preprints have not been certified by peer review, they should not be relied on to guide clinical practice or health-related behavior nor should they be reported in the news media as established information [3]. Nonetheless, the launch of preprint servers has led to the increasing use of preprints in the clinical and health science research community [4]. Although the COVID-19 pandemic highlighted the benefits of preprints, including the rapid and open communication of research findings [5-7], concerns remain that early and open public

access to preliminary medical research may harm patients or public health practices, by possibly expediting a misleading finding or faulty research that has been conducted or interpreted in error [8]. In this review, the history, current status, strengths, and weaknesses of preprints as well as current concerns regarding journal policies toward preprints are discussed.

HISTORY OF PREPRINTS

In 1961, the Information Exchange Groups was introduced by the USA's National Institutes of Health (NIH) to facilitate the distribution of preprints in the biological sciences. Until 1966, this system attracted many researchers and produced

more than 1.5 million copies of preprints, but it was then restricted because journals refused to publish articles that had been made available as a preprint [9,10].

In August 1991, Paul Ginsparg launched an electronic bulletin board intended to serve a few hundred colleagues working in a subfield of theoretical high-energy physics. Its range of topics later expanded to include physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering, systems science, and economics. It began at Los Alamos National Labs, but as its popularity grew it was relaunched as the 'arXiv' (www.arXiv.org) server, hosted by Cornell University. arXiv is considered the first and largest preprint platform [2]. Additional preprint servers followed for different academic fields, such as BioRxiv (2013, Biology; www.biorxiv.org) and medRxiv (2019, Health Science; www.medrxiv.org). A list of preprint servers and their policies and practices across platforms can be found at www.asapbio.org/preprint-servers.

Currently available preprint platforms are divided into 1) profit (such as PeerJ Preprints, Nature Precedings and F1000Research) and non-profit (such as arXiv, BioRxiv and medRxiv), 2) general (posting nearly all preprints from a wide range of disciplines, such as Authorea, Preprints, Academia, and ResearchGate) and field-specific (such as bioRxiv, medRxiv ChemRxiv and EarthArXiv), and 3) regional (such as INA-Rxiv, Frenxiv, AfricArxiv, Arabxiv).

The introduction of the overlay journal, which does not produce its own content but selects from articles on preprint servers, contributed to the growth of preprint servers both quantitatively and qualitatively [11].

CURRENT STATUS OF PREPRINTS

The number of preprint submissions has increased over the years and reached a peak in 2019 (Fig. 1). The cumulative numbers of preprint submissions are 2,185,334 at arXiv, 180,133 at bioRxiv, 100,805 at Research Square, and 48,849 at medRxiv. The cumulative numbers of preprint downloads continue to increase and, in 2019, reached 2,641,473,782 at arXiv, 144,168,644 at bioRxiv, and 63,895,520 at medRxiv (Fig. 2).

The National Library of Medicine launched 'NIH Preprint Pilot' in June 2020 with the goal of making preprints stemming from NIH-supported research accessible via PubMed Central (PMC) and PubMed. NIH Preprint Pilot consists of

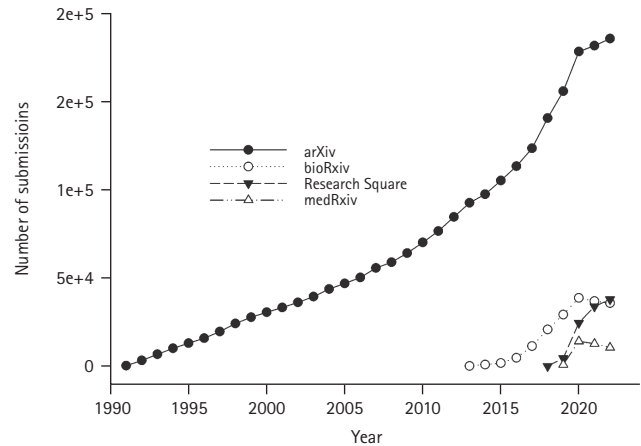


Figure 1. Number of preprint submissions categorized by individual preprint platforms.

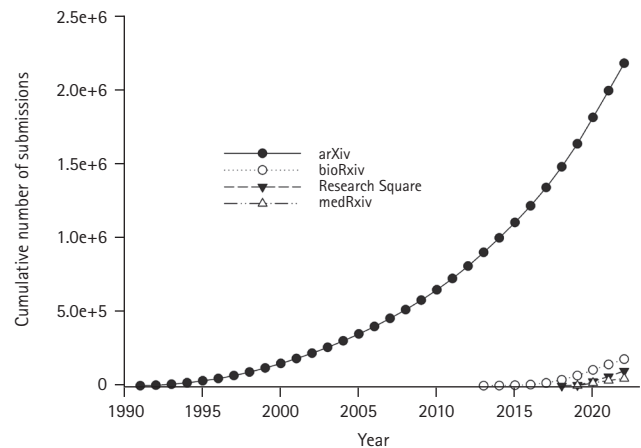


Figure 2. Cumulative number of preprint submissions categorized by individual preprint platforms.

two phases. Phase 1, which ran from June 2020 to December 2022, focused on improving the discoverability of preprints related to SARS-CoV-2 and COVID-19 research carried out with NIH support. Phase 2, which began in January 2023, covers all preprints, whether stemming from research with direct NIH support or involving an NIH-affiliated author, posted to an eligible preprint server (bioRxiv, medRxiv, arXiv and Research Square) on or after January 1, 2023. Further information on NIH Preprint Pilot is available at NIH Preprint Pilot - PMC.

As of January 31, 2023, 49,792 preprints had been released on medRxiv. This is a significant increase compared to the server's first complete month, July 2019, in which only 129 preprints were released, and ~40 times more preprints

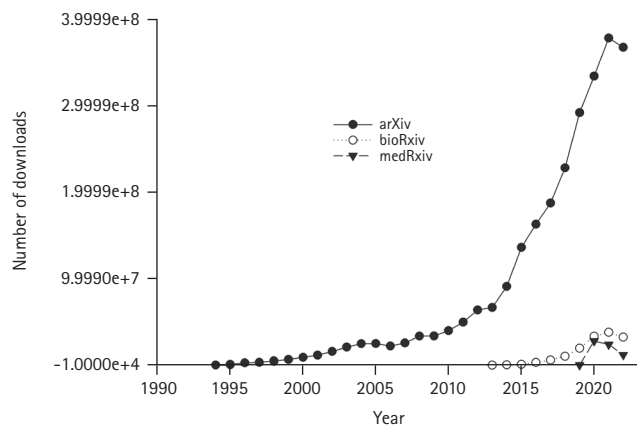


Figure 3. Number of preprint downloads categorized by individual preprint platforms.

than available in January 2020 (1,289 preprints). The number of preprint submissions to medRxiv reached a high point in 2019, followed by a peak in the number of downloads in 2020 (Fig. 3). These increases can be attributed largely to research focused on the COVID-19 pandemic.

The scientific response to the COVID-19 pandemic coincided with and promoted an unprecedented approach to research communication, one based on rapid, open-platform reporting of research results and considerable developments within preprint research literature. Both are considered essential to guide timely, evidence-based public health responses during infectious disease outbreaks and other public health emergencies [6,12].

STRENGTHS AND WEAKNESSES OF PREPRINTS

Preprint posting of non-peer-reviewed work enables rapid access to information by circumventing possible drawn-out journal submission or publication procedures. Preprint platforms have become increasingly prevalent and have also bridged the gap between academic and non-academic audiences by providing public access to research on a wide range of topics [13]. However, preprints are not peer-reviewed, which has both advantages and disadvantages. In several notable cases, this has allowed the dissemination of unsupported, faulty conclusions across several media channels. In these cases, preprints have received harsh criticism and several have since been rescinded.

CURRENT CONCERNS OF PREPRINTS

The rapid propagation of preprints has raised several issues. These are discussed in the following sections, from the positions of editorial board members, peer reviewers, and authors.

Does a journal allow a manuscript with preprints to be submitted for formal publication?

Due to the exponential expansion of preprints throughout the COVID-19 pandemic, submitting a manuscript with a preprint has become commonplace [5], which has forced journals to establish policies for those submissions. Due to limitations in its capacity for editorial service and/or a preference for avoiding chaotic preliminary situations, a journal may choose not to allow the submission of a manuscript with a preprint. Alternatively, a journal may accept the submission of manuscript with a preprint by requiring the author to report the existence of the preprint and its DOI. The preprint policies of publishing companies vary and include “can share”, “should allow,” and “encourage”.

In terms of the number of preprint DOIs, an editorial board may choose to accept preprints with only one DOI, in the case that a preprint has been posted to multiple preprint platforms. Tracking a preprint with two DOIs doubles the load of editorial services and may lead to substantial confounding situations.

Is a preprint citable as a reference in the manuscript?

Among the preprints in medRxiv, 77% (1,077 out of 1,399) were published in peer-reviewed journals within a median of 6 months after posting [14]. In a journal article that assessed how COVID-19 evidence presented in preprints changed after review, point estimate values were found to have changed by an average of 6%, with a strong correlation (0.99) between estimate values before and after review [15]. Although a substantial portion of preprints were later published as peer-reviewed journal articles, and the robustness of the evidence reported in preprints was found to be high, a preprint is preferably cited as a reference after the corresponding peer-reviewed journal article has been published, because the peer review process reduces uncertainty and substantiates the evidence reported in the preprint. It is also crucial that preprint authors disclose the preliminary

nature of the reported findings to the reader and refrain from overstating the outcomes of any study that has not undergone peer review. If a journal accepts preprint citations in journal submissions, the reference format should make clear that a preprint has been cited and the non-peer reviewed status of the source [16].

Is a double-blind review process necessary in the review of manuscripts with preprints?

In a double-blind review process, neither authors nor reviewers are aware of one another's identity. This system is used to minimize review biases, whether toward authors in the reviewer's own co-author networks or those against competing corresponding authors. For these reasons, double-blind review is a crucial component of the peer review process for publications and disciplines.

However, because many decisions are made by editors, double-blind peer review cannot completely eliminate bias. In addition, many reviewers who seek to identify the authors are often successful. Preprints in particular may make it more challenging to maintain blinding because a manuscript with a preprint discloses the identities of the authors to peer reviewers. COPE (Committee on Publication Ethics) prohibits posting pseudonymous preprints, but there are other information sources that can jeopardize blindness. For instance, randomized clinical trials registered at clinical trial registries can be accessed by peer reviewers and will reveal the authors' names. Likewise, seminar announcements, conference programs or papers, or social media posts are all searchable online [17]. These issues are not specific to preprints, but they do further complicate their management.

When a manuscript has a preprint, reviewers are also given the opportunity to comment and to provide input on it. Remarks about the study's merits and faults as well as any unrecognized research controversies may assist the reviewer in his or her task but they may also harm the review process. As a result, it is crucial that reviewers carefully and critically consider the obtained information and strive to avoid prejudging the work described in the manuscript.

How to handle changes in the evidence reported between the original preprint and the peer reviewed manuscript?

A peer reviewed, published journal article should be considered a cornerstone of evidence, even if controversies remain. Further discussion and review may lead to additional

evidence and a subsequent peer-reviewed article. Therefore, after the research in the preprint is formally published in a journal, the original preprint should be hyperlinked to the published article to reflect the change in content.

In addition, as the publisher has the copyright on a journal article published after peer review, it is better to cite a peer-reviewed journal article than a preprint.

Is it acceptable that the authors of a manuscript are different from those of a preprint?

Authorship should be limited to researchers with significant intellectual, social, and financial contributions. It should also imply responsibility and accountability for published work.

Many preprint platforms do not yet define the authorship criteria for preprints. However, some preprint platforms, including Preprint.org, define authorship in accordance with the International Committee of Medical Journal Editors (ICMJE) recommendation. Several preprint platforms also address reporting and ethical considerations, such as clinical trial registration, competing interest statement, patient/participant permission, and ethics committee approvals.

If the authors of a submitted article differ from those of the preprint, the authors must justify the change in authorship and demonstrate that it complies with ICMJE recommendations.

Is the submission of a manuscript with unintentional 'scoop' issues allowed?

A 'scoop' occurs when researchers publish their findings before a competing team working on the same issue has published theirs, or when an idea or set of results are referred to in a publication without proper attribution.

In a field that is competitive and moves quickly, a research team (team A) may choose to publish its results in preprints rather than in reputable publications because being first is more important to the team than being thorough. If another research team (team B) later submits equivalent data to a journal, what should the journal's acceptance policy be? The implementation of most research projects is time-consuming. If the work of team B is released shortly after that of team A, it is impossible for team B to have begun and completed a completely new research project within the short time span. In such cases, team A may have unintentionally scooped team B and manuscript submission by the latter may be allowed.

Being scooped is one of the worst things that can happen to a researcher. If he or she intended to submit work to a journal after posting a preprint but the results and conclusions are almost the same as those already submitted, the manuscript may be rejected by the journal editor or no longer qualify for publication in a better journal. Journals should therefore always consider potential scoop issues during the publication review process and look for and assess papers published on related topics.

How to enforce the review process for a submitted manuscript with a preprint?

Enormous interest among scholars following the posting of a preprint on a platform could result in information overload and disorderly circumstances. In addition to an experienced peer review of a submitted manuscript, its preprint history needs to be tracked and summarized by an additional, independent editor focused on the posting procedure and format rather than on the academic considerations. The findings should then be shared with peer reviewers and the journal editor.

Journals should also decide whether to consider comments posted on preprint platforms before and/or during the peer review process; if so, standard formats and mechanisms should be established as to how these will be considered.

Influence of social media and publication ethics

In 2020, a report on the use of hydroxychloroquine to treat COVID-19 was posted on medRxiv and simultaneously accepted for publication in an open-access journal. On the same day, the president of the USA publicly stated his belief in this treatment. Prescriptions of hydroxychloroquine surged 2,000% within 2 months, and US stockpiles of hydroxychloroquine reached 63 million doses within 4 months. This preprint was discussed 1,027 times on Twitter and posted 75 times on an internet portal blog. Academic societies debated the effectiveness of hydroxychloroquine, and the therapeutic efficacy was eventually shown to be extremely poor. Nonetheless, the article was cited 3,024 times according to the Web of Science, and the publishing journal's impact factor increased from 5.283 in 2020 to 15.250 in 2021, which was largely attributable to the article [18].

The resulting waste of medical resources was not due solely to the released preprint, as the improper media cover-

age of a scientific finding without adequate peer review also played a major role. This case also demonstrates that editors and authors should not use journal articles and preprints for ethically questionable purposes.

In conclusion, the policy regarding preprints is still evolving. Journals may allow the submission of a manuscript that was posted on a preprint platform. Preprints can be cited as references after peer-reviewed journal articles are published but they should only be cited alone in exceptional circumstances, without overstating the work reported in the preprint. Special considerations for dealing with scoops, authorship, publication ethics, and social influence should be provided.

REFERENCES

1. Callaway E. Preprints come to life. *Nature* 2013;503:180.
2. Ginsparg P. ArXiv at 20. *Nature* 2011;476:145-147.
3. MedRxiv. About MedRxiv [Internet]. New York (NY): MedRxiv; 2023 [cited 2023 Mar 1]. Available from: <https://www.medrxiv.org/content/about-medrxiv>.
4. Massey DS, Opare MA, Wallach JD, Ross JS, Krumholz HM. Assessment of preprint policies of top-ranked clinical journals. *JAMA Netw Open* 2020;3:e2011127.
5. Fraser N, Brierley L, Dey G, 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.
6. Brierley L. Lessons from the influx of preprints during the early COVID-19 pandemic. *Lancet Planet Health* 2021;5:e115-e117.
7. Watson C. Rise of the preprint: how rapid data sharing during COVID-19 has changed science forever. *Nat Med* 2022;28:2-5.
8. Flanagan A, Fontanarosa PB, Bauchner H. Preprints involving medical research-do the benefits outweigh the challenges? *JAMA* 2020;324:1840-1843.
9. Cobb M. The prehistory of biology preprints: a forgotten experiment from the 1960s. *PLoS Biol* 2017;15:e2003995.
10. Confrey EA. The information exchange groups experiment. *Publ Res Q* 1996;12:37-39.
11. Eysenbach G. Celebrating 20 years of open access and innovation at JMIR publications. *J Med Internet Res* 2019;21:e17578.
12. Yozwiak NL, Schaffner SF, Sabeti PC. Data sharing: make outbreak research open access. *Nature* 2015;518:477-479.

13. Fox F. The preprint dilemma: good for science, bad for the public? A discussion paper for the scientific community [Internet]. London (UK): Science Media Centre; 2018 [cited 2023 Mar 1]. Available from: <https://www.sciencemediacentre.org/the-preprint-dilemma-good-for-science-bad-for-the-public-a-discussion-paper-for-the-scientific-community/>.
14. Janda G, Khetpal V, Shi X, Ross JS, Wallach JD. Comparison of clinical study results reported in medRxiv preprints vs peer-reviewed journal articles. *JAMA Netw Open* 2022;5:e2245847.
15. Nelson L, Ye H, Schwenn A, Lee S, Arabi S, Hutchins BI. Robustness of evidence reported in preprints during peer review. *Lancet Glob Health* 2022;10:e1684-e1687.
16. JHL Editorial Team. Policy about preprints and other non-peer reviewed research. *J Hum Lact* 2021;37:17-18.
17. Lin YH, Sahker E, Shinohara K, et al. Assessment of blinding in randomized controlled trials of antidepressants for depressive disorders 2000-2020: a systematic review and meta-analysis. *EClinicalMedicine* 2022;50:101505.
18. Journal citation reports: International Journal of Antimicrobial Agents. Philadelphia (PA): Clarivate; 2021.

Received : March 2, 2023
 Revised : March 14, 2023
 Accepted : March 25, 2023

Correspondence to

Hyun Kang, M.D., Ph.D., M.P.H.
 Department of Anesthesiology and Pain Medicine, Chung-Ang University College of Medicine, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea
 Tel: +82-2-6299-2583 Fax: +82-2-6299-2585
 E-mail: roman00@naver.com
<https://orcid.org/0000-0003-2844-5880>

Acknowledgments

Hyoung-Chul Oh is the current Editor-in-Chief of *Korean Journal of Internal Medicine* and Hyung Kang is the current Editor-in-Chief of *Anesthesia and Pain Medicine*. However, they were not involved in the peer reviewer selection, evaluation, or decision process of this article. This manuscript is published simultaneously in the *Korean Journal of Internal Medicine* and *Anesthesia and Pain Medicine*.

Credit authorship contributions

Hyoung-Chul Oh: conceptualization, data curation, formal analysis, methodology, project administration, visualization, writing - original draft, writing - review & editing; Hyun Kang: conceptualization, methodology, data curation, formal analysis, writing - original draft, writing - review & editing, visualization, project administration

Conflicts of interest

The authors disclose no conflicts.

Funding

None