Preprints Involving Medical Research—Do the Benefits Outweigh the Challenges?

Annette Flanagin, RN, MA; Phil B. Fontanarosa, MD, MBA; Howard Bauchner, MD

A preprint is a complete manuscript posted to a preprint server by authors before peer review and publication in a journal. The goals of preprints are to enable authors to obtain timely feedback and comments on research before submission to a peer-reviewed journal, to claim provenance of an idea, and to facilitate and expedite dissemination of and access to research. Preprints can be amended or updated, commented on by others, and remain on the preprint server even if subsequently published in a journal. They can be cited and indexed and increasingly are given attention in the news and social media.

In clinical medicine, the ultimate aim of research is to improve patient outcomes and public health. Whether preprint posting and rapid dissemination of non–peer-reviewed reports of medical research that could have important clinical implications and consequences help achieve the goal of improving health outcomes for patients without causing harm remains uncertain.

Preprint servers, which are increasing in number, host and archive preprint manuscripts. Considered the first preprint server, arXiv was launched in 1991 for physics researchers to share scientific reports with each other before journal publication. Before that, in 1961, the US National Institutes of Health began a preprint program for sharing biological preprints, known as Information Exchange Groups, but this program was discontinued in 1967 after journals refused to consider submissions previously posted as preprints. In 2013, bioRxiv was launched for preprints in biology and the life sciences, and in 2019, medRxiv, dedicated to health sciences, began. As of September 2020, there were at least 61 public preprint servers covering many disciplines; one-third (21) of these have been launched since 2018, and an increasing number permit the posting of preprints in medicine and health.

Preprint servers are managed and supported by a range of financial models, including support from professional societies, nongovernmental organizations, foundations, and funders, and more recently, large publishers, and some servers require a fee for preprint posting.

In this issue of *JAMA*, Malički and colleagues report a cross-sectional analysis of 57 of the largest open (not funder-associated) preprint servers and identified 10 that have posted more than 500 preprints in the health sciences. The authors analyzed the policies of preprint servers that included screening before posting, submission requirements, and 18 recommendations on transparency in reporting and research integrity (eg, data sharing; addressing plagiarism, image manipulation, and correcting errors; reporting conflicts of interest, funding, and ethics approval; and guidance on authorship and reporting). Most preprint servers in the study (82%; n = 47) had some form of, albeit minimal, screening. Of the submission requirements, all servers required a specification of the scholarly scope of preprints and 54% (n = 31) required an indication of the type of study permitted. Regarding the assessment of transparency in reporting and research integrity recommendations, more servers that post preprints in the health sciences (40%-60%), compared with all servers (16%-39%), had recommendations about data sharing; plagiarism and correcting errors; and reporting conflicts of interest, funding, and ethics approval. However, very few of any preprint servers provided guidance on authorship (14%; n = 8), image manipulation (4%; n = 2), and reporting study statistics (n = 0) or study limitations (4%; n = 2) or following recommendations of the International Committee of Medical Journal Editors (9%; n = 5) or the Committee on Publication Ethics (4%; n = 2).

Also in this issue, Krumholz and coauthor-founders of the medRxiv preprint server reviewed content on the server and trends 1 year after it was launched in June 2019. The founders of medRxiv met with several members of the *JAMA* editorial staff before launching the site to solicit feedback, and many of the issues raised in that meeting were ultimately addressed by medRxiv. Krumholz et al report the posting rate of submissions after passing screening criteria, which include the following: “the manuscript is a full scientific research report (not a narrative review, commentary, or case report); the absence of obscenity, plagiarism, or patient identifiers; and confirmation by an affiliate (a member of the scientific community who voluntary screens submissions) that posting would not pose potential risk to patients or public health.”

During its first year, medRxiv received 11,164 submissions, with large increases in recent months; 9967 (89%) of these submissions passed screening and were posted. The authors do not report the reasons that 11% were rejected for posting; however, they do report that only 18 (0.002%) of these preprints have been withdrawn after posting, and 13 of those withdrawals were related to coronavirus disease 2019 (COVID-19). Krumholz et al also report on increasing numbers of posts, downloads, and views during the debut year, especially during recent months. They indicate that COVID-19 submissions comprised 73% of the total preprints posted between February and June 2020 and that 12% of COVID-19–related submissions did not meet the screening criteria and were not posted.

The authors also report that only 9% of medRxiv preprints have received comments and only 10% have been...
published in peer-reviewed journals. The time frame for some portion of these preprints may have been too short to capture eventual journal publication, but the low rate of on-site comments raises questions whether preprints are an effective mechanism for authors to obtain feedback prior to submission to a journal. The authors note that these low rates of on-site comments do not include comments on social media. One possible explanation is that the goal of speed to dissemination has become more paramount than scientific community engagement and review before journal peer review and publication. This may reflect the COVID-19 pandemic and may not be representative of preprints in general.

Benefits and Challenges of Medical and Health-Related Preprints

Clinicians have addressed the benefits and challenges of making preprints and early manuscript drafts of new research findings widely and publicly available before accuracy, reliability, and potential bias of studies have been vetted through the traditional editorial and peer review processes. A convenience sample survey of 512 researchers, librarians, publisher representatives, and many other stakeholders (including relatively small numbers of students, government and nonprofit agency representatives, clinicians, industry researchers, journalists, preprint server providers, research administrators, and funders) was conducted in June and July 2020 by ASAPbio (Accelerating Science and Publication in Biology) to assess views about the benefits and risks of preprints. The majority of respondents (66%) were researchers. More than 90% of respondents indicated the top benefits of preprints as “increasing the speed of research communication” and “being free to read.” However, 79% indicated they were concerned about “premature media coverage of preprints,” and 63% reported they were concerned about the “public sharing of information before peer review.”

In addition to having a screening process that may be more stringent than that of other preprint servers, medRxiv includes a note of caution on its home page: “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.” medRxiv has added a similar note of caution, and the arXiv physics server has added a warning specifically for its COVID-19 e-prints (ie, “they should not be relied upon without context to guide clinical practice or health-related behavior and should not be reported in news media as established information without consulting multiple experts in the field”).

In addition, each individual medRxiv preprint carries a cautionary note in the online version that it “reports new medical research that has yet to be evaluated and so should not be used to guide clinical practice.” However, the caution for the news media included on the medRxiv homepage is missing from the online version of medRxiv preprints, and the PDF versions include none of these warnings. Given that many users will access preprints directly and share them with colleagues and not via server homepages, such cautionary warnings would be best displayed on all versions of preprints, and other preprint servers should consider adopting similar cautionary notices.

A recent study by Fraser et al, released to date only as a preprint (caution advised), assessed the role of bioRxiv and medRxiv preprints during the first 4 months of the COVID-19 pandemic (January-April 2020). This study found substantially higher numbers of COVID-19-related preprints posted in the first 4 months of the pandemic compared with the total number of preprints posted about Zika virus or Ebola virus during the entire durations of those epidemics, 2015-2016 and 2014-2016, respectively. The study also found that COVID-19-related preprints received substantially higher attention in news and social media compared with non-COVID-19 preprints. Fraser et al observed that despite the warning messages provided by medRxiv and bioRxiv, “COVID-19 preprints have received unprecedented coverage on online media platforms” and suggested that “this represents a marked change in journalistic practice: pre-pandemic, bioRxiv received very little coverage in comparison to journal articles.” These authors also found weak correlations between news and social media attention to preprints and citations and concluded that because most COVID-19 preprints had not yet been published, “concerns regarding quality will persist.” Indeed, manuscripts posted on preprint servers during the COVID-19 pandemic have been widely noted and cited by various media outlets. Any notion that preprint servers are only for the scientific or medical community is incorrect.

There is a general assumption that more rapid access to information will improve patient outcomes—the ultimate goal of research in clinical medicine. However, it is quite clear that in some countries, information from social media and preprint servers has been used by politicians and physicians to advocate for specific treatments. In the US, early in the pandemic, both hydroxychloroquine and convalescent plasma were advocated by various individuals prior to any evidence suggesting benefit. This type of advocacy makes conducting randomized clinical trials more difficult and may lead to inappropriate use of some drugs and potential harm. For the majority of research in clinical medicine, posting manuscripts on preprint servers will not affect clinical care, but for studies that are likely to influence clinical care, specifically certain clinical trials and observational cohort studies, investigators should pause before posting early reports on a preprint server and consider the potential consequences.

Posting a manuscript on a preprint server before peer review and publication provides more information and is more helpful than press releases issued by investigators or companies citing a successful new treatment before peer review and publication. However, many journals have the capacity, on a limited basis, to conduct expedited editorial evaluation and peer review and to publish manuscripts in a matter of weeks. For example, recently JAMA published 3 clinical trials and a meta-analysis on the use of corticosteroids for patients with COVID-19 within 2 weeks after the last of the 4 manuscripts was submitted. None of these reports had been posted as preprints. Although peer review is not without challenges and some limitations, the process does provide an important check.
and balance on the appropriate reporting of the conduct, analysis, interpretations, and conclusions of a study.

Journal Guidance for Authors and Journalists Regarding Preprints

A recent study of 100 top-ranked clinical journals found that 86% of journals allowed preprints and 13% had a case-by-case assessment policy.\textsuperscript{19} Even before the COVID-19 pandemic, journal preprint policies and guidance for authors had been evolving. For example, in 2018, JAMA and the JAMA Network journals discouraged authors from submitting manuscripts that had been previously released to the public as preprints. In 2019, the JAMA Network journals changed the policy to a case-by-case determination: “Public dissemination of manuscripts prior to, simultaneous with, or following submission to this journal, such as posting the manuscript on preprint servers or other repositories, will necessitate making a determination of whether publication of the submitted manuscript will add meaningful new information to the medical literature or will be redundant with information already disseminated with the posting of the preprint.”\textsuperscript{20}

Accordingly, posting of a manuscript on a preprint server before submission to a JAMA Network journal does not preclude consideration of the manuscript for publication, although authors are advised and expected to provide information about any preprint postings, along with any other related manuscripts, at the time of submission of a manuscript to the journal. The journals also follow an embargo policy for submitted and accepted manuscripts that restricts news coverage until publication and advise authors to avoid promoting preprints of their submitted manuscripts in the news media and in social media before editorial evaluation and peer review and journal publication.\textsuperscript{20} To help encourage transparency, the JAMA Network also recommends that authors who cite preprints in their manuscripts indicate “preprint” in the citation in reference lists.\textsuperscript{21}

Other major journals have also modified their policies regarding preprints. For example, the New England Journal of Medicine also permits submission of manuscripts previously posted as preprints provided that authors notify the journal of any such preprint.\textsuperscript{22} In 2018, The Lancet began offering authors the option to simultaneously post preprints to its publisher-owned preprint server when submitting manuscripts. The Lancet editors report that about 30% of authors of research manuscripts submitted to the Lancet journals have opted to post preprints, but of those, “only two-thirds had all the required information (ethics approval if needed, declaration of interests, funding statement, and prospective registration for randomised controlled trials).”\textsuperscript{22} The editors also note concern about widespread press attention to preprints and the need for caution when citing such research and note a plan to “apply a more obvious watermark stating that these are preprints and not peer reviewed.”\textsuperscript{23}

In 2019, the Nature journals, including Nature Medicine, announced a move from support to encouragement of preprints and advise that authors can engage with news media about their preprint studies provided they explain that the study has not been peer reviewed and that findings could change.\textsuperscript{24} In 2020, PLOS journals amended their policy toward preprints and publication embargoes, in that manuscripts previously posted as preprints and accepted for publication remain under a news embargo.\textsuperscript{25} In doing so, PLOS reversed its previous policy of not applying an embargo to articles previously posted as preprints, citing their analysis that found that manuscripts previously posted as preprints and released to the press without embargo received significantly less media attention than a comparison group of manuscripts previously posted as preprints that were embargoed until journal publication. PLOS journals made this change to avoid a disadvantage to authors of reports previously posted as preprints and acknowledged that press embargoes are the best tool to facilitate fair and equal access to journalists and allow them time to assess new research and consult experts before dissemination to the public.\textsuperscript{25}

Guidance from press officers at the University of Leuven in Belgium concedes that preprint servers are “a goldmine for journalists looking for their next big story” and offers useful tips to researchers in dealing with the complexities of studies reported in preprints and the news media.\textsuperscript{26} Bollen and Nelissen\textsuperscript{26} advise authors to not request promotion of preprints via press offices because “press offices are not peer reviewers, and one single press release about findings that don’t hold up can cause long-lasting damage.” They also urge authors not to send their preprints to journalists to draw their attention to their work, recognizing that preprints are intended to be read by fellow researchers and that journalists and the public may not understand the difference between an unvetted preprint and a peer-reviewed article and that premature coverage may contribute to disinformation. They also advise authors if a journalist contacts them to ask about a manuscript previously posted as a preprint that is under review with a journal to ask the journal about its policy on preprints and embargoes before responding to the journalist. This guidance has long been recommended by the JAMA Network journals\textsuperscript{20} and many other medical and health journals.\textsuperscript{27}

In a recent JAMA Viewpoint, Saitz and Schwitzer\textsuperscript{28} described concerns regarding the rapid public reporting on the efficacy of hydroxychloroquine and remdesivir for treatment of patients with COVID-19 as examples of how misinformation can damage public trust in science and medicine. The authors urged caution, scrutiny, and clear and complete reporting of new study findings during the COVID-19 pandemic, including assessment of study limitations and presenting important caveats, and acknowledged that this applies to preprints and new releases as well as journal articles that are expedited to publication. Saitz and Schwitzer addressed common failures in this regard, including: “(1) a focus on single study results without the context of other studies or acknowledgment that single studies are rarely definitive; (2) overemphasis on results, particularly relative effects, without recognition of important limitations; and (3) communications based on incomplete reports of studies and reports of studies that have not been adequately reviewed.”\textsuperscript{28}

During the COVID-19 pandemic, and perhaps thereafter, investigators may continue to want their findings released and shared as rapidly as possible, but such speed to widespread
public dissemination vs sharing within a community of specialists most likely to understand the complexities of the science and concerns to public health or without rigorous editorial evaluation and peer review before publication does not come without consequences and potential for harm.\(^1\)\(^3\)\(^4\)\(^5\)\(^6\) For many investigators, preprints may be considered an initial step along the scientific dissemination and publication pathway, just as abstract, poster, and video research presentations at in-person and virtual scientific meetings have a role in the early sharing and discussion of studies among specialist communities before publication in a journal. While manuscripts previously posted as preprints may be improved following formal submission to a journal and undergoing editorial evaluation, peer review, revision, and editing, others may not be suitable for formal publication because of methodologic flaws, biases, and important limitations. Authors should share preprints during the processes of manuscript submission to journals, just as they do with study protocols and registration reports, to aid journal editors in the evaluation of the quality of the reporting of the study and prioritization for publication. Preprints and preprint servers are here to stay, but perhaps in the immediate future a more selective use of these sites may be warranted, with clinical investigators exercising caution when the focus of a study is on drugs, vaccines, or medical devices and the results of a study may directly affect treatment of patients.

REFERENCES


16. JAMA Instructions for Authors. Published September 27, 2020. https://jamanetwork.com/journals/jama/pages/instructions-for-authors


