Rise of the preprints

Nature Cancer encourages preprint sharing as a valuable means of research dissemination and scholarly communication.

When it comes to sharing research findings, preprints are nothing new. On the contrary, for some disciplines, they represent the status quo. In the physical sciences, preprint deposition in *arXiv* has been the preferred method of sharing research since the server's inception in 1991, with over 1.7 million submissions listed so far. The *Social Science Research Network* (SSRN) was originally launched in 1994 and, in the intervening decades, has expanded to include almost 1 million submissions in more than 55 disciplines. Over the years, several servers have also been launched in the life sciences, although this area's interest in preprints developed more slowly. For example, *Nature Precedings*, the life sciences preprint server started in 2007 by the Nature Publishing Group, as it was then known, numbered a few thousand submissions before it wound down operations in 2012.

Since then, a palpable change in researcher attitudes has led to an explosion of preprint platforms across disciplines and geographical regions, as well as an increase in submissions in the life sciences. Among the developments that catalyzed this change was the launch in 2013 of *bioRxiv*, the highly popular preprint server established by the Cold Spring Harbor Laboratory, and ASAPbio (Accelerating Science and Publication in biology), a scientist-driven non-profit organization that first convened in 2016 and actively promotes the adoption of preprints in the life sciences.

Indeed, preprints are an indispensable tool for fostering scholarly discourse and accelerating scientific discovery. They permit early and free dissemination of findings to a wide audience, something that is proving particularly valuable in a year when rapid access to COVID-19-related work is of huge importance. Sharing of COVID-19-related preprints has surged — at the time of writing, 2,711 preprints had been posted on *arXiv*, 2,213 on *bioRxiv*, 8,099 on *medRxiv*, 5,472 on SSRN and 3,968 on *Research Square* — to name some of the most popular platforms. The immediate and broad availability of preprints allows the assessment and discussion of scientific findings on a worldwide level, which can be especially rich when combined with social media commentary. The ability to update preprint versions allows researchers to establish a dated and citable record of their work while they revise and improve it ahead of formal peer-reviewed publication. Importantly, major funders, including the US National Institutes of Health (NIH), the European Research Council, the UK Medical Research Council, Cancer Research UK and the Wellcome Trust, to name but a few, have also voiced their support for preprint sharing in recent years. Thus, a preprint can provide evidence of an investigator's productivity and research progress before the peer-reviewed, published 'version of record' of their study is ready. Furthermore, the increased visibility provided by a preprint can help raise the professional profile of junior researchers, establish new collaborations and attract the interest of journal editors for formal peer-review and publication.

Stacked against these advantages are concerns about the posting of premature studies to establish priority of discovery or in response to public health emergencies. A major purpose of preprints is, of course, to disseminate early-stage work that can continue to be modified, but the potential to muddy the scientific waters with weak or problematic studies cannot be discounted. Many servers have been tightening their screening processes to ensure that posted preprints contain original, bona fide research that adheres to key scientific and ethical standards. As yet, however, there is no unifying set of screening practices[3] across preprint platforms. It should be noted that the ability of the wider scientific community to provide feedback can act as a real-time quality control filter for posted research. However, reversing the harm caused when incorrect and alarmist information reaches the general public can be hard. This was highlighted earlier this year by the publicity surrounding a preprint purporting that SARS-CoV-2 shared similarities with HIV that were unlikely to have occurred naturally. Rapid scientific commentary on the analyses and interpretations of the data led to the preprint's swift withdrawal, but not before it fueled theories that the SARS-CoV-2 virus had been engineered in a lab. To guard against the misinterpretation of preprint findings, servers such as *arXiv*, *bioRxiv* and *medRxiv* warn that preprints have not been peer-reviewed and thus should not be reported as established information.

A separate criticism is that the preprint literature can be hard to search. The proliferation of preprint platforms and the fact that major research databases do not index this type of non-peer-reviewed report meant that preprints were not easily discoverable. This has begun to change in recent years, with preprints from many servers becoming indexed in Google Scholar, PrePubMed, Europe PMC and OSF preprints. More recently, NCBI started a pilot scheme in which NIH-funded preprints are indexed in PubMed Central. Specific scientific communities have also established preprint assessment networks. For example, The Company of Biologists offers preLights, where biologists cover preprints of interest to the biological community. Separately, *Nature Reviews Immunology* curates the article collection COVID-19 Watch, which includes immunology- and COVID-19-related preprints highlighted by scientists from the Icahn School of Medicine at Mount Sinai, New York.

A more substantial limitation of preprints is that they often do not provide a complete version of the scientific study, given that supplemental information, such as supporting data, and details on data and code availability, is frequently omitted. This can be mitigated in part by enhancing prepost screening and improving the functionality of preprint platforms. However, preprints are not intended to replace scientific journals, which serve not only to direct the peer-review process but also to establish and uphold high standards and stringent policies for high-quality reporting and publishing of research findings. The development of stronger ties between preprint platforms and traditional publishers testifies to this synergistic relationship. For instance, *bioRxiv* and *medRxiv* permit the bidirectional transfer of studies between their servers and participating journals for formal peer review of a preprint at a journal or deposition of a manuscript as a preprint.

The support of the *Nature* journals for preprints predates the launch of *Nature Precedings* by a decade, as demonstrated by a 1997 *Nature* Editorial[1]. Our preprint policies were updated in 2019 (ref. [2]) to actively promote the use of preprints. Thus, as with all of the *Nature* journals, *Nature Cancer* encourages the posting of

preprints. Having posted a preprint would not jeopardize consideration of the study at this journal. Moreover, the version that was originally submitted to the journal may be posted as a preprint at any time during the peer-review process. Preprints may also be cited in the reference lists of *Nature Cancer* papers; however, if this is to support key methodology or central findings in the manuscript, authors may be asked to provide peer-reviewed citations or experimental evidence as further validation. Since June of this year, we have also been offering In Review, a free opt-in service that enables journal-integrated deposition of research manuscripts that have been directly submitted to the journal as preprints on the *Research Square* platform.

Such collaborative initiatives between preprint servers and traditional publishers can ensure that preprints and the peer-reviewed literature synergize to enhance the scientific record, through a dynamic process of sharing, discussing and improving research findings before and after publication. ❐

Published online: 17 November 2020
https://doi.org/10.1038/s43018-020-00151-y

References
1. *Nature* **390**, 427 (1997).
2. *Nature* **569**, 307 (2019).
3. Kirkham, J. J. et al. *bioRxiv* https://doi.org/10.1101/2020.04.27.063578 (2020).