Technical and social issues influencing the adoption of preprints in the life sciences

Naomi C Penfold[1], Jessica K Polka[1]

[1] ASAPbio, San Francisco, CA, USA

Corresponding Author:
Jessica Polka

Email address:
Jessica.polka@asapbio.org

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

Technical and social issues influencing the adoption of preprints in the life sciences

Authors: Naomi C Penfold and Jessica K Polka

Abstract

Preprints are gaining visibility in many fields. Thanks to the explosion of bioRxiv, an online server for preprints in biology, versions of manuscripts prior to the completion of journal-organized peer review are poised to become a standard component of the publishing experience in the life sciences. Here we provide an overview of current challenges facing preprints, both technical and social, and a vision for their future development, from unbundling the functions of publication to exploring different communication formats.

Unbundling the functions of publication

Science progresses only at the rate at which we can share information with one another. But as any author of a journal article can attest, formal mechanisms of scholarly communication do not always work efficiently and can be subject to biases (1–3). Peer review takes time: not merely for the reviewer to compile a thorough assessment, but also for the editor to find reviewers who are available to spend a day assessing a new manuscript on short notice. In the swiftest case, a manuscript is accepted at the first journal and the process to eventual publication may take approximately four months (4,5). However, given that many researchers continue to be evaluated based on the reputation of the journals where their work is published, authors are incentivised to 'aim high' when they select which journal to submit their manuscript to, and it can take several rounds of review (at a single or multiple journals) before the work is approved for publication. It is commonplace for a manuscript to have been submitted to at least two journals on its way to publication, and as a result the overall peer review process can take years (6).

The sooner a piece of work can be read, evaluated, and built upon, the faster science moves. And by including a greater diversity of thought in the process of science, the higher the quality of its final products. Yet, while our system of publication has superficially transitioned from physical print magazines to online websites, the mechanisms and processes of scientific communication are not much faster or more inclusive than they were in the 19th century.

Perhaps the underlying cause for this stasis is the fact that our system of evaluating scientific work—whether for deciding what to read, or to whom to award grants and jobs—relies heavily

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

on the reputation of journal titles. Experimenting with new forms of sharing science that are incompatible with publication in traditional venues carriers career risks. In addition, many open-science practices (posting lab notebooks, sharing datasets, or conducting replication studies) require significant extra effort for researchers, which is currently not well-supported with money or time. Therefore, researchers need efficient mechanisms of sharing research that align with current publishing practices, while supporting a gradual evolution towards more transparent and efficient communication practices. One small step towards a world of more transparent information exchange is to simply share manuscripts publicly at the time they are ready to send to a journal, i.e. by posting a preprint. Using preprints to separate in-depth review from the initial act of sharing can increase efficiency while requiring minimal extra work for authors and presenting science in a format that is easily recognized by readers.

Here we distill what we've learnt from our work listening to concerns about, and investigating issues surrounding, preprints. We summarize the current state of support for preprinting in the life sciences, discuss extant needs and challenges, and put forth ideas for future developments.

Why now?

Posting preprints is standard practice in many fields in physics, mathematics, computer science, economics, and other disciplines. Preprints are only now becoming widespread in the life sciences, despite a long history of sincere efforts to establish servers in biology by both public and private sectors dating back to the 1960s (6). Why have they taken off in biology only now? We suspect that at least four factors have contributed.

First, in today's digital world,the idea of composing a manuscript in real-time using collaborative editing tools only to *not* share it with the community seems increasingly anachronistic.

Second, bioRxiv was positioned effectively within the existing publishing paradigm from the start. Founded by veterans of the publishing industry, John Inglis and Richard Sever, bioRxiv quickly established partnerships with a number of journals. These journals not only agreed to consider manuscripts posted as preprints, but also established a direct submission pipeline enabling authors to submit to both with one click. Furthermore, perhaps driven by a competitive publishing environment, editors began to invite submission of manuscripts from preprint servers (discussed below). Preprints now represent an opportunity to publishers, where previous efforts to share science in this way may have been seen as a commercial threat. Direct submission arrangements and anecdotes about manuscript recruitment offered researchers confidence that the act of preprinting would not endanger their chances of journal publication. Furthermore, the ownership of bioRxiv by Cold Spring Harbor Laboratory, a credible, non-profit research institute, likely contributed to its resonance with the community of authors and readers.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

Third, many funders have since provided active support and recognition for preprints. While the NIH has been involved in preprinting through the Information Exchange Groups of the 1960s (Cobb) and Harold Varmus's 1999 eBioMed proposal (7), only recently have many funders voiced support for preprints as a mechanism for applicants and grantees to demonstrate productivity. We discuss these policies in detail below.

Fourth, Twitter created a community that provided visibility to preprints and support to their authors (8). All of the benefits of preprinting (including discussion, collaboration, visibility, and earlier disclosure) rely on active acknowledgement of preprints by the authors' community. At the early stages of any movement, supporters will be relatively far and few between, limiting their ability to support one another. Twitter has allowed preprint enthusiasts to connect with one another across institutional boundaries, meaning that even a small number of early adopters can reap the benefits of increased exposure and feedback for their work by sharing preprints with one another.

Preprints in harmony with journals

In 1966, a cabal of journal editors "outlawed" Information Exchange Groups (the NIH's photocopy and mail-based preprint exchange platform), fearing that preprints would damage their business model (6). A representative of the American Association of Immunologists wrote that "Since the preprints are complete publications, there is a real danger that they will reduce the usefulness of existing journals in the field of Immunology and may ultimately supersede them."(9) Indeed, reports that papers change little between their preprint version to the final published version have caused some to declare that preprints can be the end of the story (10). Despite the obvious irony that the article reporting this similarity added a whole section on bioRxiv during review, the more serious issue is that textual analysis may not accurately capture significant changes in meaning. And there is value in evaluation even if the manuscript stays exactly the same: peer review can provide validation as well as improvements.

Perhaps for these reasons, authors continue to use journals even in fields in which preprinting has long been common practice. For example, in physics, 73% of papers on the arXiv can be matched to an article that appears in a journal indexed by Web of Science (11). While bioRxiv is younger, the number is similar (67%, (12)), suggesting that neither archive is massively disrupting the journal business.

In fact, preprints are very much complementary to journals, and they offer several tangible benefits for editors and publishers. Preprints allow authors to receive feedback from a broader range of scientists than could be engaged in a typical peer review process. This means that the version of the paper that is ultimately accepted by the journal will have undergone more scrutiny, likely leading to a higher quality final product.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

Furthermore, preprints offer an efficient marketplace for papers. While many editors travel to conferences to invite submission of future manuscripts based on interesting presentations, preprint servers make the manuscripts themselves open to review by anyone in the world. Therefore, it is no surprise that the practice of inviting journal submissions from preprint servers seems to be widespread (13). *PLOS Genetics* has pioneered the formalization of this process with preprint editors (14) and *Proc B* has adopted the practice as well (15). Unfortunately, many such invitations may be moot since it is common practice for authors to post the preprint version concurrently with submission to a journal, a process that is facilitated by integrations in both journal and bioRxiv submission systems (16,17). In order to allow this marketplace of submission invitations to function efficiently, authors can post their preprint a few weeks before journal submission and allow their work to recruit feedback, attention and editorial invitations. Doing so could help save both authors' and editors' time along the way.

Finally, preprints relieve pressure on journals. Authors generally would like their papers to be published as soon as possible, leading some journals to promise shorter peer review turnaround times, perhaps at the expense of allowing reviewers to be as thorough as they would like to be (18). If authors can instead share a preprint immediately, they are likely to feel more comfortable waiting a bit longer for high-quality, journal-organised peer review.

Journal policies explicitly permitting or even encouraging preprinting have removed much lingering fear of rejection due to prior publication conflicts. Even some long-standing holdouts, notably Cell Press, JACS, and the American Association for Cancer Research (19) have updated their policies to be friendlier to preprints. A full list of basic journal policies on preprint archiving can be found at SHERPA/RoMEO (20), more informal lists can be found at Wikipedia (21) and detailed policies on preprint version, licensing, and media coverage policies can be found in Transpose (22).

Institutional and funder support

Preprints allow researchers to demonstrate their most recent work to prospective and current funders. It is becoming less acceptable to cite work that is "in submission" or "under review" in grant applications: where a manuscript is prepared, reviewers wish to see it and may request the applicant cites a preprinted version (23). Practically, preprints allow reviewers to judge applicants for funding or promotion by the rigor of their latest science.

In comparison to journals, university policies for the assessment of applications for hiring, promotion and tenure seem slower to change (24), but there have been bright spots for preprints. For example, in late 2016, NYU Langone Medical Center added language to their promotion & tenure guides to include preprints as a potential research output, and in early 2018, UC Davis added a "preprints" category to the their online faculty evaluation database. UT Austin, The

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

Rockefeller University, and UC Santa Cruz have all added language inviting job applications for faculty positions to submit preprints as well (25).

Perhaps the most proactive support for preprints has come from funders, who seemed poised to actively encourage the use of preprints in the life sciences. In May of 2016, the Simons Foundation Autism Research Initiative (SFARI) announced it would change its grant award letter to "strongly encourage" investigators to post preprints and that such papers would be taken into consideration in funding decisions (26). On September 1 of the same year, these concepts became integrated into the overall Simons Foundation policy, and other funders followed suit, including The Leona M. and Harry B. Helmsley Charitable Trust, EMBO long-term fellowships and Young Investigator program, Human Frontiers Science Program, MRC, Wellcome Trust, HHMI, Cancer Research UK, BBSRC, UKRI Future Leaders Fellowship program, CNRS and the European Research Council (27).

Perhaps the most influential funder policy has been NIH's guide notice NOT-OD-17-050, which clarifies the NIH's position on preprints and other interim research products (28). "The NIH encourages investigators to use interim research products, such as preprints, to speed the dissemination and enhance the rigor of their work...Interim research products can be cited anywhere other research products are cited." A notable exception, however, is in the use of preprints in post-submission materials (29), which are intended to accommodate events outside the control of the investigators.

Some private funders have gone beyond encouraging preprints to requiring them. Barring privacy concerns, the Chan Zuckerberg Initiative states a commitment to posting preprints prior to peer review (30). As part of Wellcome's updated open access policy, researchers working on fields of public health relevance will be required to preprint at the time of journal submission from 2020 (31).

As with all policies, their existence does not ensure they will be enacted. Funders also must develop mechanisms to monitor grantee reaction and compliance. The emergence of technological infrastructure (for example, links between preprints and published papers, metadata about funding sources, and submission and posting dates), as well as continued dialogue between researchers and funders, is key to enabling these policies.

Technical issues

At present, preprint servers lack the technological instructure that could help them to realize their full potential. Addressing such challenges could make a large impact on how preprints are used and discovered.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

For example, authors who have previously read a preprint often wish to quickly find out how it has changed upon the posting or publication of a subsequent version. Currently, neither preprint servers nor journals present a summary of the changes made. Some users already make version notes when posting a revised manuscript to bioRxiv; making this more standard practice might involve enabling authors to submit a short piece of text to journals as well, similar to a conflict of interest disclosure or author CRediT declaration. Once this is complete, it would be natural for journals to provide a link back to the preprint version, which would present a more complete picture of how a manuscript evolved over time. Some journals already provide this backwards link — including Nature Machine Intelligence, Plant Direct, and PLOS One (33–35). Preprints could also be better supported by reference managers with features that would allow users to link preprints to later versions (whether revised preprints or a final journal version) and receive updates when subsequent versions are available online.

Change is needed in search tools, too. For example, preprints could also be linked from PubMed and PubMed Central. (Note that this is effectively being done for papers in *F1000 Research* and associated platforms such as *Wellcome Open Research*. Once these papers pass peer review, they appear on PubMed Central along with their date-stamped first version.) This helps to establish a record of what work was done when, irrespective of delays imposed by the peer review process, which is key to determining priority of discovery. Europe PMC has already implemented links between the preprint and published version of the same piece of work, though improved metadata could facilitate further search and tool development (32).

Beyond the basic metadata about a preprint, open access to the data detailing interactions with each preprint would enable innovation around how the latest science is discussed. For a recent effort to understand Twitter interactions with and downloads of preprints posted on bioRxiv, content metadata was derived by scraping the bioRxiv website (12). In the absence of an official bioRxiv application programming interface (API), these authors and others have developed their own tools (including an API, command line tool and Python wrapper) to source and interact with bioRxiv content data.

Addressing the technical issues detailed above may help more people find and interact with preprints. As we will discuss in the next section, the low discoverability and perceived legitimacy of preprints is at the root of several more complex social problems.

Social issues

Today, preprinting is treated as standard practice — or at least supported to a considerable degree — in some life science communities, such as neuroscience, bioinformatics, evolutionary biology and ecology (Abdill & Blekhman, 2019; see also subject-specific initiatives like PeerCommunityIn (https://peercommunityin.org/) and servers hosted at OSF Preprints (https://osf.io/preprints/). Other subject areas have less experience and thus may have lower

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

228 awareness of the actual benefits and issues. In addition to new servers (33–35), several new
229 research categories have been added to bioRxiv in recent years — clinical trials, epidemiology,
230 paleontology, pathology, and pharmacology and toxicology (note their absence in older literature
231 (12,36)). This freshness demands and enables considered discussion of important issues so that
232 the most beneficial practices surrounding preprinting can be cemented as cultural norms. A
233 recent consultation highlighted that researchers were often unable to cite case studies of the
234 benefits of preprints (8), and so continued productive adoption may require increasing the
235 number and visibility of shared real-life experiences with preprints (such as those at
236 wesupportpreprints.wordpress.com (37)).
237

May 2018	• "PLOS and bioRxiv announces a partnership where PLOS authors can also opt to share their articles on bioRxiv." (10.31222/osf.io/796tu) • Crossref reports that preprints are growing at 10x the rate of articles
June 2018	• The Lancet launches a preprint platform on SSRN • African scientists launch their own preprint repository, AfricArxiv
July 2018	• Europe PMC announces it will now index preprints • PLOS announces they link to the preprint from the published article page
August 2018	• Journal of the American Chemical Society (JACS) permits manuscript submissions that have been preprinted on arXiv, bioRxiv and ChemRxiv • ERC indicates 2019 plans to highlight that preprints can be cited in applications (PDF)
September 2018	• PKP and SciELO announce development of open source Preprint Server system to interoperate with OJS and other SciELO journal systems
November 2018	• Wellcome Trust will require grantees to preprint research where there is a significant public health benefit from January 2020 (now updated from be from January 2021)
December 2018	• ICMJE adds recommendations for medical publishing conduct with respect to preprints • The Israel Science Foundation announces the upcoming launch of ISF Open Research as an open peer review platform for research funded by its programs
January 2019	• EcoEvoRxiv launches as a preprint server for ecology and evolutionary biology

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

February 2019	• bioRxiv starts rollout of full-text HTML conversion for all preprints • AMRC Open Research officially launches as an open peer review platform for research funded by AMRC member charities
April 2019	• Beilstein Journals post first preprint in their preprint server for organic chemistry and nanotechnology
May 2019	• PLOS has posted 2,500 preprints to bioRxiv through author opt-in upon submission in the first year of the PLOS-bioRxiv preprint-posting partnership • ORCID adds preprint as a 'work type' and supports the addition of works using arXiv IDs, enabling authors to document their own preprints in their record • Springer Nature unifies preprint policies on licensing, citation, and media coverage "to encourage preprint sharing"
June 2019	• ResearchSquare's pre-publication platform, In Review, has expanded and now covers 33 journals and platforms (it launched in 2018 with four BMC journals) • MedRxiv, a collaboration between CSHL, Yale, and *The BMJ*, launches.
August 2019	• Open Access India and COS launch IndiaRxiv

Table 1. Developments in preprinting across biomedical and life sciences since May 2018, adapted from ASAPbio (38) and additional web search. For developments before May 2018, refer to Tennant et al, 2018 (35).

Licensing

While open access to scholarly literature has been discussed for decades, its original meaning has been diluted. The Budapest Open Access Initiative describes "free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself." (39). Today, the majority of articles on PubMed Central, while "free" to read, are not actually open for reuse. Articles not in the OA subset cannot be downloaded in bulk, restricting access to text and data mining (40). Even if that bulk file were available, their licenses do not permit reuse.

Because authors are directly in control of licenses on preprints, they have an opportunity to create a more open corpus of literature. However, most authors on bioRxiv are choosing restrictive licenses (41) amid widespread confusion about what they mean and a misconception that journals prohibit the use of certain licenses for preprints (42). In reality, we are aware of

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

only a single publisher with this policy, IOP, and it has limited coverage in the life sciences. In contrast, an influential funder, the NIH, has recommended the use of CC BY (28). More education and guidance for authors is needed, for example within the preprint submission process itself. Ideally, however, co-authors would have an informed discussion about the license to choose for their preprint *before* submission.

Permitted versions

The term "preprint" can describe many different versions of a manuscript, ranging from drafts shared for feedback well before journal submission to manuscripts ready to be accepted by a journal. However, journals differ in their policies regarding which versions of manuscripts under consideration may be posted, with some of them prohibiting the posting of preprints after initial submission. These policies may be rationalized by a sense of journal ownership of the peer review process, but in fact they prevent scientists from sharing improvements drawn from diverse sources—their own additional experiments and analysis, feedback colleagues with whom the manuscript was privately shared, comments on the preprint server itself, and input from social media and preprint-specific feedback platforms (including preLights, PREreview, biOverlay and PeerCommunityIn). Adding to the confusion, preprint servers differ in their own policies for manuscript deposit; in many disciplines (canonically, arXiv) preprint servers also host postprints, or versions of manuscripts after journal acceptance. In the life sciences, PubMed Central, complemented by institutional repositories, fulfills this need, and bioRxiv hosts only preprints, not postprints. However, other repositories can host biology postprints, for example OSF Preprints.

Scooping

A common fear cited as a barrier to preprinting is "getting scooped." Researchers may feel this has happened when a competing research group publishes highly related work without crediting (i.e. fairly citing and discussing) their own preprint. As a consequence, their work receives less attention and recognition, and if the work is still unpublished, can mean publication in a "lower" journal.

It stands to reason that scooping fears are most acute when the stakes are high and careers are on the line. However, fears about scooping – and the secrecy that accompanies them – cannot be neatly divided by generations because it's rare for a group of co-authors to be homogenous in years of experience.

Fear of scooping impacts not only researchers' willingness to share preprints at all, but also whether they are willing to share auxiliary materials that are normally shared as a condition of journal publication. For example, communities have yet to come to consensus on whether authors should be obligated to share reagents or strains after posting a preprint. In a future world

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

where preprinting is universally regarded as a respected disclosure, ethical standards of disclosure should match those associated with journal articles.

Curation and evaluation

As the production of scientific outputs continues to accelerate, both as a result of a growing number of researchers and their increasing willingness to share, we will need new ways of dealing with information overload. While an overabundance of publications may feel like a 21st century problem, thinkers since Seneca have lamented the overabundance of information, and scholars have developed tools to help organize and filter it (43).

Currently, readers report finding preprints by searching for keywords. They also report being alerted to interesting work on Twitter. The first strategy is directed by subject area, but not interest, and the second by interest, but not subject area. Ultimately, we will need more efficient ways to combine both search criteria in a single stream, in much the way that journal title is presently used (rightly or wrongly) to help parse search results in PubMed. Rxivist is one such tool that marries current interest and subject area (12), and we are collecting more curation projects at reimaginereview.asapbio.org. We believe that this emerging space will become an essential component of the preprint ecosystem.

Curation of interesting or highly-respected preprints can also improve their usefulness in evaluating scientists for jobs and grants. While journal name (and Impact Factor) are flawed proxies for judging the quality of a work (44), they save reviewer time by quickly communicating information about a paper's selection process. Such proxies are not essential in the late stages of an evaluation process when candidates have been whittled down to a short list and reading their full outputs is a manageable task. However, the process of shortlisting candidates requires more time-efficient indicators of research quality than reading the content itself. Shortly after publication, such indicators may include the level of authors' transparency and openness, endorsements from peers, and assessments of creativity. In the longer term, established reproducibility or replicability and impact on science or society can also be assessed (45). Preprints offer the opportunity to evaluate researchers based on their most recent work, but candidates may need to accompany them with indicators that distill community reactions in the short-term, such as downloads, citation counts, constructive preprint comments, and other endorsements. Despite existing limitations, multiple reports suggest preprints are already helping early-career researchers to secure their next research position (37). Improved practices for filtering, curating, and signaling interest in preprints can further promote this phenomenon.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

The future of preprints

Who's at the table?

The growing adoption of preprints in biology is being largely driven by researchers in North America and Europe: of the top 100 institutional affiliations ranked by number of preprints posted to bioRxiv until December 2018, only 6 are located outside these regions (12). Researchers who feel comfortable posting a preprint are likely to be those who do not feel so threatened by the 'scooping' concerns identified above as to not preprint.

As a mechanism for sharing and consuming the latest science irrespective of social hierarchies, we must ensure that preprint infrastructures and social mechanisms develop with issues of diversity, equity and decolonialisation of scholarship in mind (46,47). Who can contribute to the preprinted literature? Who benefits from posting a preprint? Who can read, consume and use information in preprints? As preprinting continues to grow in biology, we must bake these questions into every discussion.

Reflecting on the 'scooping' concerns listed above, we should consider how preprints could offer appropriate recognition and support for creators of openly shared work. Indeed, some researchers report only being rewarded with funding and jobs when they are authors of (high-impact) journal articles, and not for reuse of their open datasets (48). Therefore, it can be difficult to argue that the researchers producing the primary datasets should share these openly, let alone rapidly with a preprint. This issue does not relate to the development of new tools and methods — in this case, researchers report valuing the immediate usage, testing, and feedback that preprinting these resources provides.

Once work is shared openly, it is important to address how widely it is seen. Twitter is a major driver of attention on preprints, and social connections between preprint authors and readers raise visibility in the absence of dissemination through journals. Thus the visibility of preprints is strongly influenced by the authors' existing network 'connectedness' and therefore is vulnerable to the same under-representation issues we face elsewhere in science. There have been several initiatives to increase the visibility of under-represented scientists (including VanguardStem and 500 Women Scientists (49,50)): following suit, SBotLite is a new Twitter bot that retweets preprints posted by female first authors in the hope of raising their visibility (51). Ensuring that the dissemination of preprints does not mimic or perpetuate diversity issues in STEM requires continued investment in initiatives to counteract and mitigate existing attention biases.

Beyond the article

Some have expressed concern at the roughly 35% of preprints that do not go on to be published in a journal, believing that these preprints must be of low quality (52). Alternatively, these

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

outputs could reflect work never destined for a journal that would have otherwise not been shared or work that the authors have chosen not to submit to a journal. Such products include negative results, preliminary findings, methods and protocols, and short reports from projects that could not be completed (for example, because funds or a training period ran out). All of these products are valuable, and all could be in principle posted on a preprint server. In fact, bioRxiv contains specialized sections for contradictory and confirmatory work, though they are seldom used. As of the time of writing, the Contradictory and Confirmatory Results sections together make up less than 3% of the articles on bioRxiv.

These low usage rates suggest that our current incentive system does not sufficiently reward investments of energy spent writing up contradictory or confirmatory findings in the format of a journal article. Some of this effort, for example carefully assembling a methods section, is necessary to reproduce the work, and must not be compromised. But some of the work needed to write up an article describing such findings, like putting the work in context with an introduction or interpreting the findings in a discussion, is less useful to specialized readers, who are the likely audience for contradictory or confirmatory findings anyway. In fact, those readers do not need the element of a narrative (often constructed post-facto) that ties together figures in a traditional paper. In these cases, a single figure (or even a micropublication, defined for these purposes as a statement with attribution (53)) would suffice.

There is presently an expectation that all products appearing on preprint servers are more or less complete articles. This helps to promote an image of the preprint server as a destination for high-quality work and helps to facilitate some very positive behaviors, such as the soliticiation of submissions by journal editors. However, this norm reinforces a culture in which research is shared relatively late in the process and also feeds some behaviors that are less desirable, such as counting the number of papers on a CV as a measure of productivity without assessing their contents. While this practice makes little sense, it is a real concern, as evidenced by the fact that the Medical Research Council worded its preprint policy to discourage researchers from “salami slicing” their preprints into many smaller units for the purpose of gaming the system by gaining a higher publication count (54). It is not useful to science for researchers to split one story into multiple parts purely to game the evaluation system; however, given the deeply complex and technical interdisciplinary work that is now often combined into a single 1500-word article, there is clear value in ensuring each finding is comprehensively described. If posting single figures or smaller increments of work were to become standard practice, all research results could be communicated faster and with adequate methodological description to ensure reproducibility. Those ultimately destined for a journal could be assembled into an article when the authors felt ready. Another benefit of micropublications is that they enable peer review on a more atomic level. In an environment in which papers result from the collaboration of many different specialized experts, there may be situations in which no two or three reviewers have sufficient expertise to cover every figure panel.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

Despite the apparent benefits of micropublications and preprints, both technical and social innovation is required to address open questions. Namely, how can science be shared in varying orders of detail, complexity, and review status over time, from first observation of a result to acceptance of a generalized finding into broader understanding? Which research outputs (data, code, methods) are useful to embed in a narrative article? For which of these outputs is subsequent filtration and curation valuable? Ultimately, where it is most useful to invest resources in coordinated peer review, journal production processes, and dissemination of findings to non-specialist communities? Regardless of when or how preprints fit into this picture, we should strive to ensure that research integrity is rewarded, discovery is accelerated, and the publication process is more inclusive and equitable.

Funding disclosure

The authors received no specific funding for this work.
ASAPbio is funded by:

- The Leona M.and Harry B. Helmsley Charitable Trust
- The Wellcome Trust (UK)
- Chan Zuckerberg Initiative (US)
- Howard Hughes Medical Institute (US)
- Simons Foundation (US)
- Medical Research Council (UK)
- Canadian Institutes of Health Research (CA)

Acknowledgements

We thank Anna Hatch, Daniela Saderi, Kristen Ratan, Maria Levchenko, and Michael Parkin for helpful feedback on the manuscript.

References

1. Tregenza T. Gender bias in the refereeing process? Trends Ecol Evol. 2002 Aug 1;17(8):349–50.
2. Shen YA, Webster JM, Shoda Y, Fine I. Persistent Underrepresentation of Women's Science in High Profile Journals. bioRxiv. 2018 Mar 8;275362.
3. Murray D, Siler K, Larivière V, Chan WM, Collings AM, Raymond J, et al. Gender and international diversity improves equity in peer review. bioRxiv. 2019 Apr 11;400515.
4. Royle S. Waiting to happen II: Publication lag times [Internet]. quantixed. 2015 [cited 2019

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

May 31]. Available from: https://quantixed.org/2015/03/16/waiting-to-happen-ii-publication-lag-times/
5. Himmelstein D. Publication delays at PLOS and 3,475 other journals. Satoshi Village [Internet]. 2015 Jun 29 [cited 2019 May 31]; Available from: https://blog.dhimmel.com/plos-and-publishing-delays/#journals_wrapper
6. Cobb M. The prehistory of biology preprints: A forgotten experiment from the 1960s. PLOS Biol. 2017 Nov 16;15(11):e2003995.
7. Kling R, Spector LB, Fortuna J. The real stakes of virtual publishing: The transformation of E-Biomed into PubMed central. J Am Soc Inf Sci Technol. 2004;55(2):127–48.
8. Chiarelli A, Johnson R, Pinfield S, Richens E. Practices, drivers and impediments in the use of preprints (Phase 1 report) [Internet]. 2019 May 1 [cited 2019 May 31]. Available from: https://zenodo.org/record/2654832#.XPEaXNNKg_U
9. Dray S. Information Exchange Group No. 5. Science. 1966 Aug 12;153(3737):694–5.
10. Klein M, Broadwell P, Farb SE, Grappone T. Comparing Published Scientific Journal Articles to Their Pre-Print Versions -- Extended Version. Int J Digit Libr [Internet]. 2018 Feb 5 [cited 2019 May 31]; Available from: http://arxiv.org/abs/1803.09701
11. Larivière V, Sugimoto CR, Macaluso B, Milojević S, Cronin B, Thelwall M. arXiv E-prints and the journal of record: An analysis of roles and relationships. J Assoc Inf Sci Technol. 2014;65(6):1157–69.
12. Abdill RJ, Blekhman R. Tracking the popularity and outcomes of all bioRxiv preprints. bioRxiv. 2019 Jan 13;515643.
13. Slavov N. Why I love preprints [Internet]. Slavov Lab. 2017 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20180828220418/web.northeastern.edu/slavovlab/blog/2017/09/28/biomedical-preprints-benefits/
14. Barsh GS, Bergman CM, Brown CD, Singh ND, Copenhaver GP. Bringing PLOS Genetics Editors to Preprint Servers. PLOS Genet. 2016 Dec 1;12(12):e1006448.
15. Barrett Spencer C. H. Proceedings B 2017: the year in review. Proc R Soc B Biol Sci. 2018 Jan 10;285(1870):20172553.
16. Frequently Asked Questions (FAQ) [Internet]. bioRxiv. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821203507/https://www.biorxiv.org/about/FAQ
17. Advancing the sharing of research results for the life sciences [Internet]. bioRxiv. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821203542/https://www.biorxiv.org/about-biorxiv
18. Anderson K. The Tincture of Time -Should Journals Return to Slower Publishing Practices? [Internet]. The Scholarly Kitchen. 2017 [cited 2019 May 31]. Available from: https://scholarlykitchen.sspnet.org/2017/03/28/the-tincture-of-time-should-journals-return-to-slower-publishing-practices/
19. Greene C. Why we preprint. [Internet]. Casey Greene. 2015 [cited 2019 May 31]. Available from: https://medium.com/@greenescientist/why-we-preprint-fb3bfbcdf4ff
20. SHERPA/RoMEO - Search - Publisher copyright policies & self-archiving [Internet]. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821204015/http://www.sherpa.ac.uk/romeo/search.php
21. List of academic journals by preprint policy - Wikipedia [Internet]. [cited 2019 Aug 21]. Available from: https://web.archive.org/save/https://en.wikipedia.org/wiki/List_of_academic_journals_by_p

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

reprint_policy
22. Transpose [Internet]. [cited 2019 Aug 21]. Available from: http://web.archive.org/web/20190821204315/https://transpose-publishing.github.io/#/
23. Bishop D. Tweet [Internet]. Twitter. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821204445/https://twitter.com/deevybee/status/1127186891533639681
24. McKiernan EC, Schimanski LA, Nieves CM, Matthias L, Niles MT, Alperin JP. Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations [Internet]. PeerJ Inc.; 2019 Apr [cited 2019 May 31]. Report No.: e27638v2. Available from: https://peerj.com/preprints/27638
25. University policies and statements on hiring, promotion, and journal license negotiation [Internet]. ASAPbio. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821210111/https://asapbio.org/university-policies
26. Spiro J. SFARI supports preprints for the life sciences [Internet]. SFARI. 2016 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821212320/https://www.sfari.org/2016/05/20/sfari-supports-preprints-for-the-life-sciences/
27. Funder policies [Internet]. ASAPbio. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821212524/https://asapbio.org/funder-policies
28. NOT-OD-17-050: Reporting Preprints and Other Interim Research Products [Internet]. NIH. 2017 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821212644/https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-050.html
29. After My Application is Submitted, Can I Include a Copy or Citation of a Preprint as Post-submission Materials? [Internet]. NIH Extramural Nexus. 2018 [cited 2019 Aug 21]. Available from: https://nexus.od.nih.gov/all/2018/03/02/post-submission-materials-can-i-include-a-copy-or-citation-of-a-preprint/
30. Science Initiative Privacy Principles [Internet]. Chan Zuckerberg Initiative. 2018 [cited 2019 Aug 21]. Available from: http://web.archive.org/web/20190821212941/https://chanzuckerberg.com/privacy/science-privacy-principles/
31. Wellcome updates open access policy to align with cOAlition S [Internet]. Wellcome. 2019 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821213118/https://wellcome.ac.uk/news/wellcome-updates-open-access-policy-align-coalition-s
32. Levchenko M. Preprints in Europe PMC: reducing friction for discoverability [Internet]. Europe PMC Blog. 2018 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821213831/http://blog.europepmc.org/2018/07/preprints.html
33. medRxiv [Internet]. The Yoda Project. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821214815/https://yoda.yale.edu/medrxiv
34. Narock TW, Goldstein E. Quantifying the growth of preprint services hosted by the Center for Open Science [Internet]. Open Science Framework; 2019 Apr [cited 2019 May 31]. Available from: https://osf.io/5fk6c
35. Tennant J, Bauin S, James S, Kant J. The evolving preprint landscape: Introductory report for the Knowledge Exchange working group on preprints. 2018 May 17 [cited 2019 May

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

31]; Available from: https://osf.io/preprints/metaarxiv/796tu/
36. Inglis J, Sever R. bioRxiv: a progress report [Internet]. ASAPbio. 2016 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821214557/https://asapbio.org/biorxiv
37. We support preprints… – Let's Accelerate Scientific Publishing In The Life Sciences! [Internet]. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821215234/https://wesupportpreprints.wordpress.com/
38. News from around the web [Internet]. ASAPbio. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821215354/https://asapbio.org/news-from-around-the-web
39. Read the Budapest Open Access Initiative [Internet]. Budapest Open Access Initiative. 2002 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821215658/https://www.budapestopenaccessinitiative.org/read
40. Open Access Subset [Internet]. PubMed Central. [cited 2019 Aug 21]. Available from: https://web.archive.org/save/https://www.ncbi.nlm.nih.gov/pmc/tools/openftlist/
41. Himmelstein D. The licensing of bioRxiv preprints. Satoshi Village [Internet]. 2016 Dec 5 [cited 2019 May 31]; Available from: https://blog.dhimmel.com/biorxiv-licenses/
42. Preprint licensing survey – ASAPbio [Internet]. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821220137/https://asapbio.org/licensing-survey
43. Blair A. Too much to know : managing scholarly information before the modern age [Internet]. Yale University Press; 2010 [cited 2019 Aug 21]. Available from: https://www.worldcat.org/title/too-much-to-know-managing-scholarly-information-before-the-modern-age/oclc/601347978
44. Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ. 1997 Feb 15;314(7079):498–502.
45. Moher D, Naudet F, Cristea IA, Miedema F, Ioannidis JPA, Goodman SN. Assessing scientists for hiring, promotion, and tenure. PLOS Biol. 2018 Mar 29;16(3):e2004089.
46. Okune A. Decolonizing scholarly data and publishing infrastructures [Internet]. Africa at LSE. 2019 [cited 2019 May 31]. Available from: https://web.archive.org/web/20190821221035/https://blaogs.lse.ac.uk/africaatlse/2019/05/29/decolonizing-scholarly-data-and-publishing-infrastructures/
47. Albornoz D, Chan L. Power and inequality in open science discourses [Internet]. Diversity, Equity and Inclusion panel/OpenCon, 2017; [cited 2019 May 30]. Available from: https://periodicos.ufpe.br/revistas/IRIS/article/download/238912/30639
48. Chambers C, Morey C, Open Science Working Group, School of Psychology/CUBRIC, Cardiff University. 2017 Survey on Open Research Practices [Internet]. 03:52 PM [cited 2019 Aug 22]. Available from: https://mfr.osf.io/render?url=https://osf.io/dmfke/?action=download%26mode=render
49. VanguardSTEM [Internet]. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821222639/https://www.vanguardstem.com/
50. 500 Women Scientists [Internet]. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821222728/https://500womenscientists.org/
51. sBotLite (@sbotlite) [Internet]. Twitter. [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821222840/https://twitter.com/sbotlite
52. Anderson K. Comment on Two New Initiatives at eLife To Start the Eisen Era [Internet].

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019

2019 [cited 2019 Aug 21]. Available from: https://web.archive.org/web/20190821222256/https://scholarlykitchen.sspnet.org/2019/08/15/two-new-initiatives-at-elife-to-start-the-eisen-era/#comment-83759

53. Clark T, Ciccarese P, Goble C. Micropublications: a semantic model for claims, evidence, arguments and annotations in biomedical communications. J Biomed Semant [Internet]. 2014 Jul 4 [cited 2019 Aug 21];5(28). Available from: https://jbiomedsem.biomedcentral.com/articles/10.1186/2041-1480-5-28

54. The MRC supports preprints - News and features [Internet]. Medical Research Council. 2017 [cited 2019 Aug 21]. Available from: https://web.archive.org/save/https://mrc.ukri.org/news/browse/the-mrc-supports-preprints/

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27954v1 | CC BY 4.0 Open Access | rec: 10 Sep 2019, publ: 10 Sep 2019