On the value of preprints: an early career researcher perspective

Sarvenaz Sarabipour*, Humberto J. Debat, Edward Emmott, Steven J. Burgess, Benjamin Schwessinger, Zach Hensel

1Institute for Computational Medicine, Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, USA
2Center of Agronomic Research, National Institute of Agricultural Technology (IPAVE-CIAP-INTA), Córdoba, Argentina
3Department of Bioengineering, Northeastern University, Boston, Massachusetts, USA
4Carl R Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Illinois, USA
5Research School of Biology, The Australian National University, Acton 2601, ACT, Australia
6Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal

*# email: ssarabi2@jhu.edu, zach.hensel@itqb.unl.pt

Abstract

Peer-reviewed journal publication is the main means for academic researchers in the life sciences to create a permanent, public record of their work. These publications are also the de facto currency for career progress, with a strong link between journal brand recognition and perceived value. The current peer-review process can lead to long delays between submission and publication, with cycles of rejection, revision and resubmission causing redundant peer review. This situation creates unique challenges for early career researchers (ECRs), who rely heavily on timely publication of their work to gain recognition for their efforts. ECRs face changes in the academic landscape including the increased interdisciplinarity of life sciences research, expansion of the researcher population and consequent shifts in employer and funding demands. The publication of preprints, publicly available scientific manuscripts posted on dedicated preprint servers prior to journal managed peer-review, can play a key role in addressing these ECR challenges. Preprinting benefits include rapid dissemination of academic work, open access, establishing priority or concurrence, receiving feedback and facilitating collaborations. While there is a growing appreciation for and adoption of preprints, a minority of all articles in life sciences and medicine are preprinted. The current low rate of preprint submissions in life sciences and ECR concerns regarding preprinting needs to be addressed. We provide a perspective from an interdisciplinary group of early career researchers on the value of preprints.
and advocate the wide adoption of preprints to advance knowledge and facilitate career development.

Introduction

The outputs of scientific research are varied, in the form of research articles, reviews, commentaries, perspectives, theory manuscripts, methods, data, reagents, plant and animal models, computational models, patents, drugs, vaccines, software, and highly trained researchers. Researchers are primarily evaluated on their record of peer-reviewed publications in traditional journals and the perceived value of the journal where the work is published. This process has well documented limitations (1-5), which provide acute challenges for early career researchers (ECRs): graduate trainees, postdoctoral fellows and junior group leaders, who rely heavily on timely dissemination of their work to gain feedback and recognition for their efforts. Preprints are one mechanism to address some of these limitations. Preprints are online, publicly available (open access) scientific manuscripts posted by authors on dedicated servers prior to peer review and publication in an academic journal. Most preprints in the life sciences are deposited concurrently with submission to a journal, yet some authors may chose preprint deposition as the sole way of communicating their manuscript. Most of these manuscripts are of high quality and are screened to contain appropriate content for the respective preprint server. Preprint servers make work immediately available to researchers as they do not perform peer review prior to dissemination. Two of the largest preprint servers are arXiv (comprised of scientific papers in the fields of mathematics, physics, astronomy, electrical engineering, computer science, quantitative biology, statistics, and quantitative finance) and bioRxiv (repository for the biological sciences/life sciences/biomedical sciences). There are now over 1.3 million preprints on arXiv and 34,000 preprints on bioRxiv, the latter representing the work of over 150,000 researchers from 104 countries. In addition, greater than 60% of bioRxiv articles have subsequently been published in more than 1500 journals. Facing an evolving landscape for publication and evaluation of research outputs, ECRs in the life sciences must decide how to use preprints for their work. Preprint servers in the life sciences have different scopes in terms of content, subject area, language, and geographic origin of the deposited work: multiple subjects (PeerJ preprints and biorXiv), specific subjects (AgriXiv, PaleoXiv, PsyArXiv, ChemRxiv, EarthArXiv, EngrXiv, SportRxiv) and continent or language specific (AfricArxiv, IndiArxiv, Arabixiv). MedRxiv will soon focus on medicine and health sciences, which has shown the slowest uptake of preprints in the life sciences with some leading medical journals not accepting submissions of preprinted manuscripts. The adoption of preprinting as an academic practice has grown exponentially in recent years and today approximately 1–2% of articles listed in Pubmed were initially submitted as preprints (6, 7). The increasing number of biosciences preprints (7, 8), reflects a realization that preprints can ameliorate systemic issues in the existing journal based peer-review system that disproportionately impact ECRs. We discuss the many ways in which
ECRs benefit from depositing their manuscripts on preprint servers accelerating science communication and career progression.

1. **Preprints accelerate science communication that facilitates ECR career progression.** In the current scientific publishing system journals are the gatekeepers of knowledge, defining when and where manuscripts get published and who can obtain access. Publication of manuscripts in journals can take months to years leading to significant delays in communicating work (2, 3, 9). Preprints empower authors to decide when their work is ready to be shared with the scientific community. This removes delays caused by journal publication; peer-review turnaround, editorial decision making, publisher’s response times, the length of production process and re-submission cycles of rejected manuscripts. The timescale of ECR training stages are often short. Hence the protracted duration of traditional journal publishing negatively impacts ECRs seeking funding, promotion and hiring. By posting a citable preprint, ECRs can stake a claim to the performed research. Encouragingly, a number of funding institutions including the US National Institutes of Health and UK Medical Research Council take preprints into consideration in job and funding applications (10-13). This alleviates the delays in journal publication allowing researchers merit be judged on the quality of their work rather than where it is published (the San Francisco Declaration on Research Assessment) (14).

2. **Preprints can help ECRs accelerate and optimize training time and quality, research design, and funds.** The earlier we know about research performed by peers, the earlier we can incorporate this information into our own research. Early access to knowledge and data can save months to years of ECR research and training time, reduce costs and encourage risk taking. This means working in a more informed and efficient way, with a lower likelihood that our work is redundant with something being prepared for publication. Preprints make cutting edge results and methods available to ECRs on short-term postdoctoral fellowships or starting new laboratories that previously would have only been available to close colleagues prior to publication following peer review (15). For example, in biophysics and fluorescence microscopy fields preprinted methods are being used well in advance of the peer-reviewed publication in sample labeling(16, 17), instrument design (18, 19), and image analysis (20).

3. **Preprints can allow ECRs with limited funds to publish their findings.** The cost of publishing articles in journals is often multiple thousands of dollars which can be prohibitive for ECRs and researchers from low-income economies with limited funds. Per-paper processing costs of preprints are low, typically less than $50, since they bear
few editorial or administrative burdens associated with peer review; it is typically possible to cover costs of running a preprint server without article processing charges. Preprinting is a low cost means of providing open-access to work, so that outputs are available to any researcher in the world, irrespective of whether countries and institution can afford journal publication or subscription fees.

4. **Preprints in public health and medical research can boost ECR research.** Preprinting is increasing in many areas of the life sciences, but uptake in medical fields has been slower. Timely circulation of results has accelerated public health research during infectious disease outbreaks by allowing quick identification of mechanisms of disease transmission (21, 22). Restrictions on data sharing (with appropriate considerations for patient privacy and other ethical concerns) or postponing release of results until after journal peer-review have impeded research progress (21-23). Funders such as the Wellcome Trust and Gates Foundation have recognized this and now require researchers to preprint work with urgent public health research implications (24) and subsidize publication using the post-publication peer review platform from F1000 Research.

Expedited sharing of results in physiology and epidemiology as preprints can dramatically accelerate ECR research in interdisciplinary fields as well (25) as it has done in physical sciences (26, 27). In systems biology and systems medicine, preprints and open access data can provide biochemical, biophysical and physiological parameters which are key to development of complex multiscale computational models of human health and disease (28). In the absence of open, abundant, diverse and timely availability of research results, efficient use of computational models that link molecular networks to cells, organs and organ systems has been slow and challenging (29). Accelerated release of biological results and methods for data integration will promptly inform evaluation of higher resolution predictive computational models of human pathologies, boosting ECR research concerning personalized diagnostics and therapies (28, 29).

Similarly medical research is in need of immediate open innovation via open dissemination of results as manuscripts and data. Clinical trials are multi-million-dollar, years-long efforts with critically important and time-sensitive research outputs. Yet more than 70% of clinical trials deposited to the US national libraries of medicine have no associated results article. Preprints can be coupled to clinical trials databases such as (http://clinicaltrials.gov) to inform researchers in advance of journal publication, accelerating communication among basic and translational scientists, clinicians and physicians. Archiving preprints that describe methods and parameters used in clinical trials will inform the design of other trials (30). A data trial project has aimed to make this happen (31) leading to initiation of a new preprint server, MedRxiv (32), is a significant step in increasing transparency and building a sustainable culture of curating,
archiving and efficiently sharing of results via preprints in public health and medical research (33).

5. **Preprints offer community review, commenting and feedback that improves ECR manuscripts.** A typical life sciences manuscript receives feedback from two or more peer reviewers before publication. In many cases authors ask for feedback from their lab and colleagues at their university, but there is no wider round of commenting until after publication. With a preprint, other researchers can discover the work sooner, potentially pointing out critical flaws or errors, suggesting new studies or data that strengthen the argument, or even recommending a collaboration that will lead to improvements in final journal publication as preprints are not the final form of a research paper for most authors (11, 34).

6. **Preprints help ECRs form collaborations, share or receive data.** Knowledge from early communication of findings informs on the state of the field, decisions as to which lab to join (35), where to request reagents and which new techniques to adopt in a timely manner (34). The open-access policies of preprint servers facilitates this communication, with the added benefit of encouraging collaboration and informal discussion, a feature often unavailable at traditional publishing platforms.

7. **Preprints commenting provides opportunities for ECRs to develop their reviewer skills.** Public commenting on articles posted to preprint servers is uncommon (6, 36), although more frequent than commenting on journal articles. Feedback can also occur through email, and social media platforms such as twitter. In future, open peer review on preprints by researchers at all career stages will benefit the community by facilitating discussion and collaboration among laboratories. Commenting on preprints by ECRs is an opportunity to sharpen their reviewing skills and to have a voice in academic publishing. Preprint platforms, principal investigators and funding agencies can support preprint servers and implement methods to incentivise researchers to review and comment on preprints (37). Platforms such as PREreview, Peeriodicals, Peercommunity, Prelights and biOverlay have arisen to facilitate voluntary preprint-focused blogging and peer-review (38). We strongly encourage ECRs to adopt the practice of reviewing preprints and publishing their reports. A frequent concern raised about preprint review is that it increases strain on an already overstressed peer-review system. However, preprint peer review can increase efficiency in the publication process: (1) editors can identify possible peer reviewers from those who comment on preprints (including those outside the traditional pool of reviewers), (2) preprint peer reviews can be forwarded to journals along with submissions, and (3) journals can solicit submissions from authors of preprints...
with reviews/comments demonstrating that the work is rigorous. Looking at our respective interdisciplinary fields in computational modeling and systems biology, biophysics, biochemistry, plant sciences and virology, we see strong, dynamic research cultures where preprints and journal articles complement each other. With the exception of a few journals (1), the journal peer-review process remains largely opaque and confidential. Open preprint peer review can compliment journal-solicited peer review, strengthening the peer-review system and making it more efficient.

8. **Preprints speed up publication and can eliminate redundant peer reviews to make ECRs more efficient.** With the current review system, researchers experience long median peer-review times of 5-6 months on all types of submissions (2, 3) and it can often take much longer to publish research following the first submission to a journal (9). Only 20% of scientists perform 69–94% of the all journal-solicited peer reviews culminating to 63.4 million review hours, 15 million of which are spent re-reviewing rejected papers (3). These hours are spent at the expense of mentorship, research and teaching every year. The opinions of a handful of reviewers do not necessarily represent the diversity of perspectives in the scientific community (3, 39). Further, peer review does not guarantee reproducibility, with most retractions in biomedical journals being prompted by the readership performing post-publication review (40). The increasingly interdisciplinary nature of biomedical research also requires a wider range of expert reviewers, putting more strain on the peer-review system. In addition to preprint comments informing editorial decisions, preprints can accelerate the peer-review process because; (1) researchers can begin to respond to preprint comments before journal-solicited reviews are received, (2) researchers can submit higher quality articles to journals after getting feedback from preprint readers, and (3) if open preprint peer review were to become common practice reviewing of the same article could be avoided. We argue that the functions of journals to curate and evaluate research will be strengthened by effective utilization of preprints and open preprint peer reviews (41). In future, we envision that such an open preprint peer review ecosystem will benefit ECRs and the scientific community as a whole by accelerating the publication process and increase constructive feedback of one’s work.

9. **Preprints increase ECR visibility and networking.** It is critical for ECRs to be recognizable and to build a good reputation. Networking is vital for this recognition, and can increase the potential impact of ECR publications. It can also be a valuable way of receiving career development support, through peer-support from ECRs or connections with mentors. Access to networking is often inhibited by the realities of ECR life today, i.e. having the time, travel authorization and funds to travel to events to learn of cutting
edge research and to present their own work. Posting a preprint leads to a significant increase in both scientometric data (Altmetric) attention scores (which capture mentions on social media, online, and in the news) and citations of the published paper when authors had posted the work first as a preprint (42). Much discussion of preprints happens on social media, giving ECRs a new avenue to build their professional network.

10. **Preprints helps ECRs perform corrections via revisions.** In a climate in which many journals are reluctant to update manuscripts except in the case of retraction or corrections, it can cost authors thousands of dollars to publish corrections (2). A large number of peer-reviewed papers are retracted annually (43) that in some cases could be corrected instead to address errors in the original publication. Most preprint servers including bioRxiv give ECRs the platform to provide manuscript corrections and datasets as new preprints voluntarily and in a timely manner while keeping the older records as vital history of a research project. It also permits submission of supporting materials and resources that exceed journal page limits. Versioning of manuscripts by authors to narrate research progress (44) can be a tool to increase the likelihood that articles are accurate and reproducible.

11. **Publishing ALL research findings and conditions in preprints can benefit ECRs.** Negative results, which are excluded data, unreported measures and conditions (45, 46) are an important output of research (47), sharing negative results is a time saving source of knowledge for ECRs. Despite the cost and critical implications of biomedical research and clinical trials, most are not subjected to independent renalysis (which would require deposition of all data including negative results). Yet sharing of these results is uncommon and these are hard to publish findings as their inclusion in manuscripts is often discouraged by journal editors and peer reviewers (48). Novelty is often a prime criterion for being published in a journal. Thus, it can also be difficult for authors to publish null findings or replication studies. Writing separate manuscripts dedicated to negative results takes up substantial researcher time. A number of journals have emerged that are solely dedicated to publishing negative results (46), but the number of submissions to these journals have remained low resulting in one such journal shutting down (45). Preprints eliminate this problem by providing visibility to a much broader range of findings benefiting ECRs, where ECRs can include negative results accompanying the main research paper.
Figure 1. Preprints influence many facets of the scholarly landscape. Preprints are an asset for Early Career Researchers. Preprints support a vibrant research culture and impact research decisions in multiple areas of the academic endeavor. The value of preprints for the biomedical workforce and biomedical research enterprise is currently underutilized.

ECR concerns on preprinting

A number of researchers including ECRs have voiced concerns about preprinting. We review these concerns, noting that they need to be balanced against the benefits of preprints.

Preprinting leads to scooping. There is the potential that another lab with more resources could accelerate publication for competing work or repeat an interesting experiment from a preprint and publish it before the preprint’s authors do so (49). Yet several journals now provide scoop protection for preprints and acknowledge the importance of being second (50-52). Multiple independent labs reaching the same conclusion around the same time is a sign of reproducibility and soundness of the finding and should only be supported, not penalized by any scientist, journal or funding agency. Indeed preprints can be seen as a timestamp as they are posted
publicly with a Digital Object Identifier (DOI), becoming a permanent part of the scholarly record that should be referenced in journal publications. We encourage all journals to explicitly state in instructions to authors that preprints are citable. Furthermore, this concern of scooping predates life sciences preprints. We note that in many fields, it takes years to conceive ideas, perform and finalize projects for publication.

Preprinting prevents publication. Most life science journals now accept submissions of preprinted manuscripts (53). Yet, some journals have not adapted preprint-friendly policies and some publishers have confusing and self-contradictory policies. For example, one publisher’s preprint policy states that “Authors can share their preprint anywhere at any time,” including updating preprints after peer review, while the policy for one journal published by this publisher states that “we do not support posting of revisions that respond to editorial input and peer review.” Clarity from publishers and mandates from funders on preprints will reduce ECR uncertainty. We recommend that journals choose one of a few easy-to-understand preprint policies following the models used for open-access publications and for open-access publication licenses. For example:

- Preprint α (PPα): Preprint OK; updates to preprint after review OK; preprint automatically linked from journal article
- PPβ: Preprint OK; updates to preprint after review OK
- PPγ: Preprint OK; reviewer/editor comments cannot be used to improve preprint
- PPδ: No preprinted manuscripts accepted

For journals that have embraced preprints, to increase the preprinting frequency, journals can implement a policy where submitted manuscripts can be automatically sent to bioRxiv (or equivalent). The reverse of this policy is accepting direct submissions from preprint servers, currently implemented by many journals (PLOS journals and eLife and others). The international cancer genome consortia and the 4D nucleome project require presubmission to bioRxiv prior to journal submission. The rapid adoption of open access policies shows that funders are influential in shaping scientific publishing (54). Funders such as CZI, Gates and Wellcome are leading the way in mandating preprints of funded work (54, 55); we hope public funding agencies will follow their lead.

Preprints have low visibility. ECRs need to receive recognition for their work, and a common impression is that preprints are transiently recognized. This may be particularly true for publications reporting negative results which are also unappreciated in journals. In response to this we argue that the online conversation around preprints is already more robust than that around journal articles (42), and that even if it were the case, any publishing option will benefit ECRs who need to prove productivity over a short period of time. We note that bioRxiv already receives over 4 million visits per month, and that the increased rate of preprinting and the development of tools to make it easier to identify relevant preprints such as prepubmed.org mean that visibility will likely increase.
Conclusions

Preprints are already benefiting ECRs and life scientists at large, but we argue that they are underutilized and can be used in new ways to aid ECR development and increase the efficiency of scientific research. Preprints empower trainees and amplify their voices, improving their graduate, postdoctoral and early faculty experience by allowing others to learn about their work and helping ECRs form a professional network that can provide feedback, support and opportunities. We urge trainees to embrace preprints and hope that principal investigators and senior researchers encourage posting and reviewing of articles on preprint servers. The open access feature means that, preprints can raise public awareness of health and medical research in all countries, especially in developing nations where researchers struggle to gain institutional funds to publish, read and subscribe to scientific journals. Preprints also support an exciting and stimulating research culture. A strong and prevalent culture of preprinting can significantly reduce the negative impacts of the current publishing system on early career researcher work and life. It is time for our research communities and more broadly the biomedical research enterprise to embrace preprints to their full potential.

Acknowledgments

We thank the editors and reviewers for valuable comments on this manuscript. We also thank Dr. Lenny Teytelman for helpful comments on an earlier version of this manuscript.

References


33. Oakden-Rayner L, Beam A, Palmer L. Medical journals should embrace preprints to address the reproducibility crisis. International Journal of Epidemiology 2018;1-3
45. Teytelman L. Reevaluating the Quest for Negative Results. Science Editor. 2018;4(1).
49. https://blushgreengrassatafridayafternoon.wordpress.com/2015/01/07/is-being-scooped-the-flip-side-of-a-pre-print-it-was-for-us/.