

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Copyright, Fair Use, Scholarly Communication,
etc.

Libraries at University of Nebraska-Lincoln

2019

The Need for Speed: How quickly do preprints become published articles?

Rachel Herbert

Kate Gasson

Alex Ponsford

Follow this and additional works at: <https://digitalcommons.unl.edu/scholcom>



Part of the [Intellectual Property Law Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Copyright, Fair Use, Scholarly Communication, etc. by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



The Need for Speed

How quickly do preprints become published articles?

Highlights

bioRxiv's option for authors to submit biology preprints directly to journals typically speeds up publication of articles by nearly 2 weeks—an advantage for authors who are keen to get their work published quickly

Authors approach 'preprinting' in different ways: while most post a preprint before submitting to a journal, just over a third of preprints were submitted to and accepted by a journal before the preprint appeared on bioRxiv

Biology preprints with just one version are published 2 weeks more quickly than those with multiple versions

Introduction

How quickly do preprints become published articles? What benefits are there for authors considering posting a preprint and for those people who determine journal preprint policy? That's what we aim to learn in this study, by discovering a little more about the timing of submission of preprints and the relationship between submission and acceptance of manuscripts.

The dataset: 8,711 bioRxiv preprints

- posted between 2013 and 2017
- all matched in bioRxiv to a published journal article
- publication dates retrieved from CrossRef
- journal title information retrieved from Scopus
- additional publication data for select journals directly retrieved from journal websites

Preprints in scholarly communication

ASAPBio defines preprints as “unpublished draft[s] of a research paper” (Inglis & Sever, 2019) and in recent years, there has been a sharp rise in the number of preprint servers, the variety of research areas they serve (Rawlinson & Bloom, 2019; OSF Preprints), and the sheer number of preprints posted (PrePubMed, accessed 2019). This expansion and proliferation of digital activity around new scholarly works has given rise to examinations of the nature of ‘preprinting’—from the subject areas experiencing the most growth in preprint numbers—such as the life sciences, psychology and the social sciences (Narock & Goldstein, 2019)—to the positive correlation found between preprint download activity and the Journal Impact Factor of the journal in which the final published paper appears (Abdill & Blehman, 2019).

For this study, we considered 8,711 preprints on bioRxiv, the preprint server for papers on biology. The preprints were posted between 2013 and 2017 and had to be matched in bioRxiv to a published journal article. Additional data on publication dates were taken from CrossRef and select individual journals. bioRxiv's advanced functionality, combined

with the fact that it launched in the last 10 years and publishes preprints in subject areas that have only quite recently, but enthusiastically, started engaging in this form of scholarly communication, makes it a fascinating and useful way to look at trends in behaviors around preprints.

From preprint to publication

To consider the basic overall timeline first, the median amount of time it takes for a preprint to be published is 160 days (Figure 1): that's from the date a preprint is first published to the

publication date (from CrossRef). Our findings here are consistent with those reported in other similar studies: 134 days (Inglis & Sever, 2016) and, later, 166 days (Abdill & Blehman, 2019). However, the publication date of an article is as much linked to a journal's publishing schedule as it is to the readiness of an article, so this data could be hiding some other trends from view. The version history of these preprints, and the relationship between bioRxiv and journals are also relevant factors influencing publication timelines.

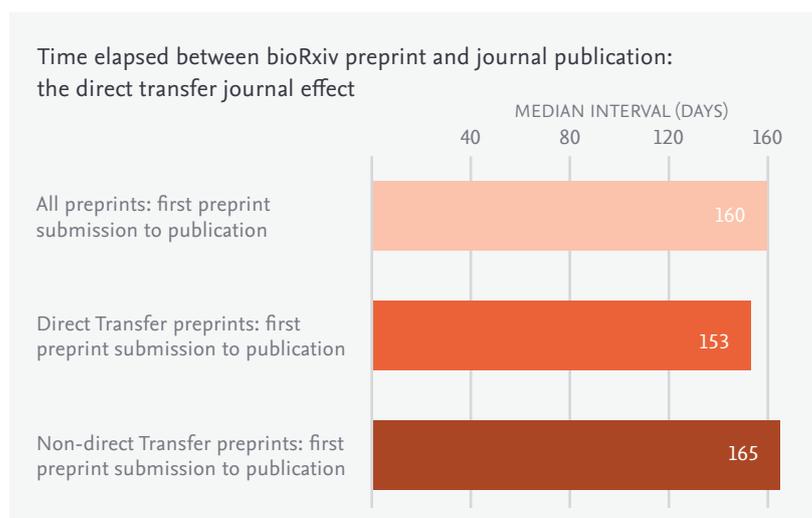


Figure 1: median interval between date of bioRxiv preprint and date of publication in a journal, comparing preprints in direct and non-direct transfer journals
Preprints in direct transfer journals: n=3,438 | Preprints in non-direct transfer journals: n=5,251
Note: 22 preprints were missing the necessary journal-level data for this analysis

Speeding up submissions

bioRxiv has a feature that is likely to play a part in the speed of publication: some journals allow direct manuscript submission via the platform. When uploading a preprint, these partner (or ‘direct transfer’) journals enable authors to directly submit a manuscript, without having to visit a different website or fill out new forms (bioRxiv.org, accessed June 2019). If authors submitting to these journals were playing Monopoly, they’d be advancing to ‘Go’ and collecting \$200 right away.

Submission to one of the 160+ direct transfer journals does not guarantee publication, and publication in a direct transfer journal does not necessarily signify that the corresponding author actually took advantage of the direct submission function. Nevertheless, preprints that were published in direct transfer journals were published more quickly than those in other journals: just under 2 weeks more quickly (Figure 1). So, when bioRxiv states that authors who use this option to submit a preprint directly to a journal “save time”... they really mean it.

Count of versions of preprints	Preprint count	Share of total count of preprints
1	5,872	67%
2	1,991	23%
3	611	7%
4	237	3%
TOTAL	8,711	100%

Table 1: preprints by the count of versions uploaded to bioRxiv

Preprints upon preprints

bioRxiv also allows users to update their preprints and so authors sometimes post multiple versions of manuscripts. The preprint webpage shows the history of each preprint and indicates the date that each version was posted. Most of the preprints in our study (67%) had just one version available across the five-year period (Table 1). Of the 2,839 preprints with multiple versions, the majority (70%) have two versions posted and 611 preprints have three versions.

Some interesting outliers reveal the range of multiple preprint version activity: one preprint had no fewer than 19 versions posted across an approximately 9-month period; another had an impressive 2 years, 28 days between the first and latest available versions.

When authors replace preprints with new versions, it could be a signal that various forms of peer review or other feedback are occurring. It’s possible that preprints are uploaded and then edited as the research study or discussion continues, or as a result of feedback from colleagues or other contacts, other bioRxiv users, and/or journal peer reviewers.

Whatever drives the need to update preprints, there is a benefit of sorts: the time from latest available preprint to publication is over three weeks shorter than that between first preprint submission and publication (Figure 2). But when we add in time spent working on additional versions (7.7 weeks median average)—overall, preprints with just one version on bioRxiv are published fastest. They are published about two weeks sooner than those with multiple versions.

Do researchers simply need to get their written work right first time?! Seemingly, uploading a single version of a preprint is the quickest ‘route’ to journal publication, after all. Well, maybe in an ideal world. But this analysis still doesn’t yet give us the full picture. We need more information about the point of submission to a journal to truly understand the potential advantages of iterating a manuscript on bioRxiv before publication.

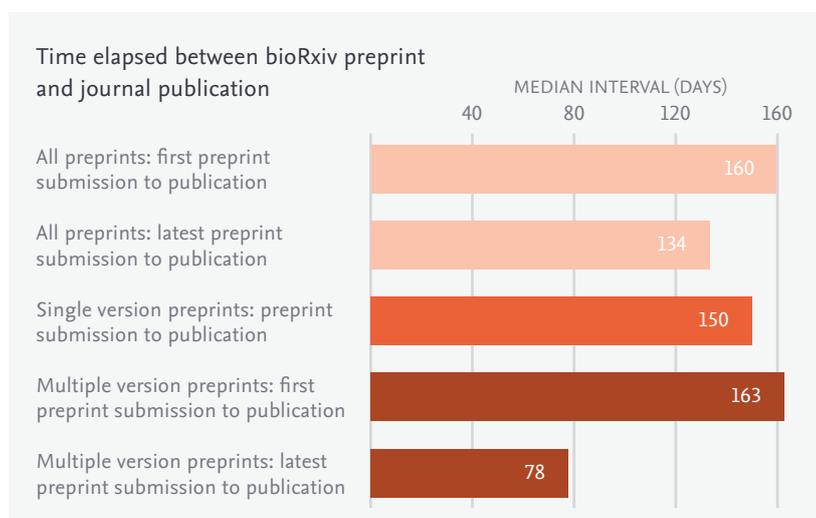


Figure 2: median interval between date of bioRxiv preprint and date of publication in a journal
All preprints: n=8,711 | Single version preprints: n=5,872 | Multiple version preprints: n=2,839

What do publication dates hide?

For a fuller picture, we need to look beyond article publication dates. As any submitting author knows, the time taken from submission to publication can be... lengthy. The availability of reviewers and the speed at which they work, the number of rounds and complexity of the reviews, and the journal's publishing schedule all play a part in constructing the timeline. So if we want to understand the benefits for authors in terms of publication speed, we need more information—in particular, dates that manuscripts were submitted and accepted. In a helpful move toward greater transparency, journals now typically publish those dates on their websites.

Eight journals publishing bioRxiv preprints were selected for this additional analysis: the four direct transfer journals and four non-direct transfer journals with the highest count of preprints (Table 2). For the 2,468 preprints across these journals, the dates the manuscript was Received, Accepted and Published were retrieved from publisher websites. We reviewed the publishing speed trends for these eight journals across 2018, and found no consistent differences in

Journal	Count of (preprint) publications	bioRxiv direct transfer journal?
Scientific Reports	557	✗
Plos One	388	✓
eLife	327	✓
PNAS	292	✓
Nature Communications	283	✗
Bioinformatics	262	✗
PLOS Computational Biology	234	✓
Nucleic Acids Research	149	✗
TOTAL	2,492	—

Table 2: the four largest direct transfer and four largest non-direct transfer journals and their count of preprint publications

the time from 'manuscript received' to 'article published' for the two groups of papers.

As shown in Figure 3, we found the following:

- Just over half (55.4%) of preprints were submitted to bioRxiv *before* they were received by a journal
- 38.6% are submitted to a journal *before* being posted to bioRxiv
- The remaining 6.0% were submitted on the *same day* as they were received by a journal
- Almost all preprints (95.5%) are posted on bioRxiv *before* being accepted by a journal

There's no doubt, then, that preprints are just that—versions of articles released before publication (in line with bioRxiv's policy)—but authors vary with respect to how far in advance of submission to a journal they make their preprint available.

Unsurprisingly, more preprints appear on bioRxiv before being received by non-direct transfer journals (61.3% of preprints in non-direct transfer journals) than by direct transfer journals (49.6% of preprints in direct transfer journals). That speaks to the (slightly) longer wait we might expect as authors select and then submit to their journal of choice rather

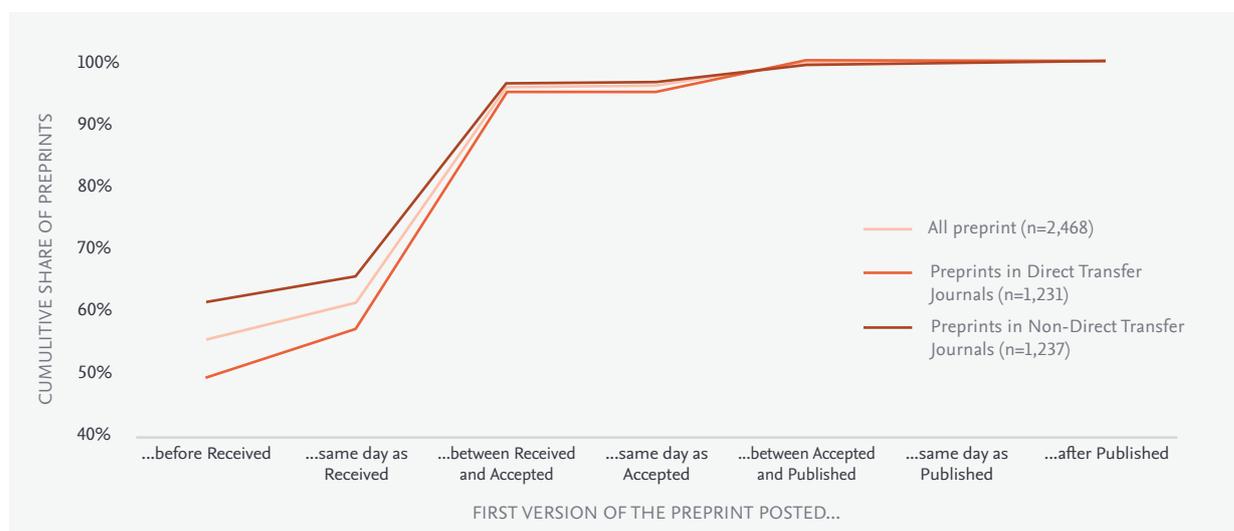


Figure 3: the timing of the first submission of preprints in relation to submission to a journal for preprints in all journals, direct transfer journals, and non-direct transfer journals.

than using the direct submission route. More preprints also appear on bioRxiv between being received and accepted for publication for direct transfer journals—and the average time between preprint posting and manuscript received is just a single day (median interval, [Figure 4](#)).

Recommendations

We conclude that bioRxiv should highlight all of the advantages of direct transfer journals. Not only do authors save time by filling out just one set of virtual paperwork when submitting preprints in these cases, they're also likely to speed up publication of their article. Although taking this route may not be the deciding factor when it comes to selecting a journal, arming authors with this insight can only help them navigate the process.

We also find that the ability to update preprints on bioRxiv offers an advantage. Even if the majority of authors don't update their preprint, they can make use of that functionality in whatever way suits them. Authors can adjust preprints as they receive feedback from colleagues and their network, and/or update them post

peer-review. In a world that keeps toying with post-publication peer review for articles, this option is something to be considered by journals and publishers as well.

That leads us to aspects of preprint behavior that are not yet understood. The findings in this study suggest that a form of review of preprints is occurring that drives the authors to update and replace their preprint on bioRxiv. Comparing versions of preprints and reaching out to authors to ask what drives the changes and edits will help improve this understanding.

Further, this study leaves the authors with a fascinating and, as yet, unanswered question: do the various effects of creating and posting a preprint on a platform such as bioRxiv increase the speed and success of peer review?

No doubt there will be a range of answers to these questions, but as preprints continue to thrive and expand to new subject areas and domains (e.g. Barry, 2018), understanding what behaviors and actions drive, and are driven by them will accomplish several things. It will

inform the research community about the advantages and disadvantages of preprints (e.g. Polka, 2017), inform those building the various preprint platforms about how to best serve submitting authors, and guide publishers and journal editors in their decisions on preprint policies (Teixeira da Silva & Dobránski, 2019).

Across the board, preprints sent to direct transfer journals do tend to be accepted more quickly than those sent elsewhere. This difference of 10 days is likely to be attractive to submitting authors ([Figure 4](#)).

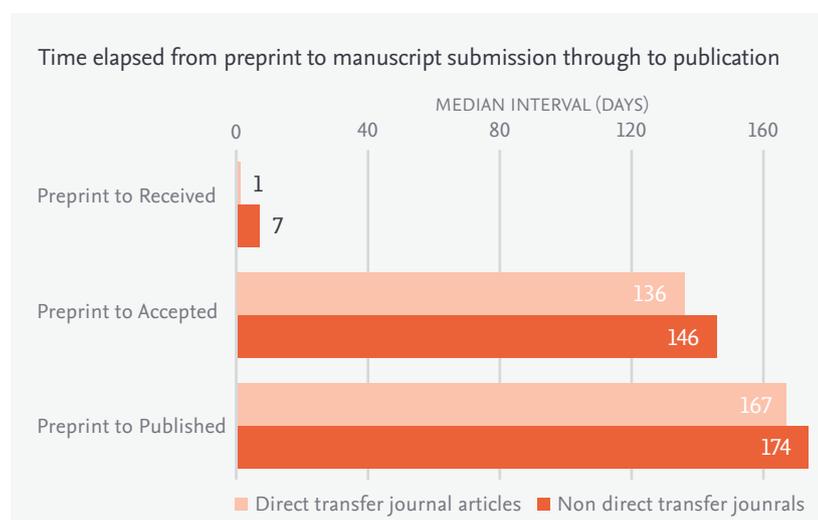


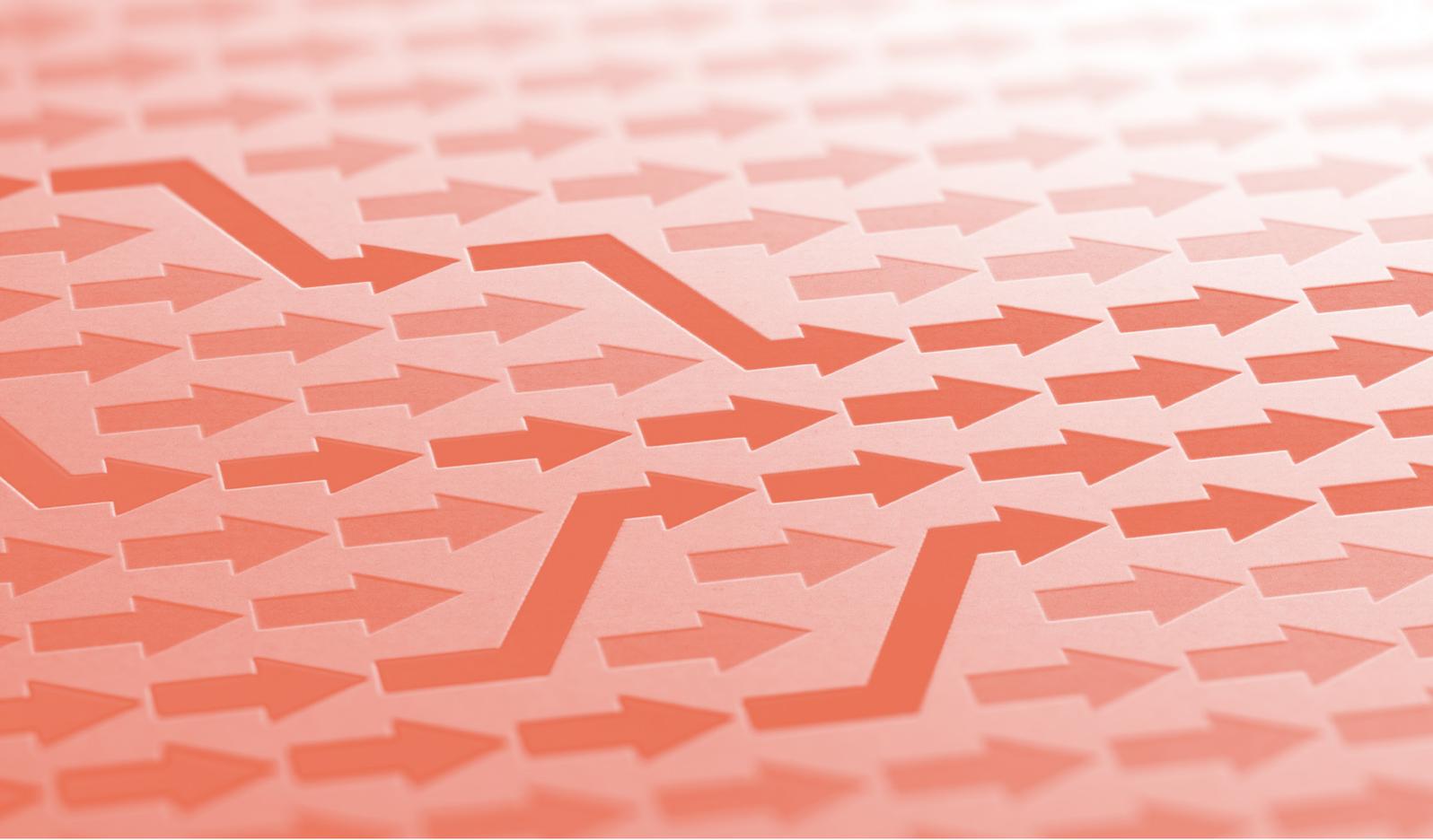
Figure 4: median interval (days) between submission of a preprint on bioRxiv and the date received, accepted and published for eight select journals; comparing four direct transfer journals to four non-direct transfer journals

Preprints in direct transfer journals with received and accepted dates: n=1,231

Preprints in direct transfer journals with published dates: n=1,241

Preprints in non-direct transfer journals with received and accepted dates: n=1,234

Preprints in non-direct transfer journals with published dates: n=1,251



Method

Using bioRxiv data, 9,122 preprints with publication dates ranging from 2013 and 2017 have been analyzed. bioRxiv has matched each of these to a published article digital object identifier (DOI). Using the DOI as a match, source titles were obtained from Scopus, and publication dates were obtained from CrossRef data; this data was available for 8,711 preprints. Additional data was retrieved from publisher sites to obtain Received, Accepted and Published dates for select journals. These journals had the highest count of publications with bioRxiv preprints and included four journals that offer a direct transfer between publication and bioRxiv, and four journals that do not currently offer that service:

- Direct transfer journals: *PLoS ONE*, *eLife*, *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, *PLoS Computational Biology*
- Non-direct transfer journals: *Scientific Reports*, *Nature Communications*, *Bioinformatics*, *Nucleic Acids Research*

These eight journals published 2,492 preprints, of which additional publication date information was available for 2,468 preprints.

References

- Abdill R.J. & Blekhman, R. (2019) Tracking the popularity and outcomes of all bioRxiv preprints. *eLife*, 8, e45133. <https://doi.org/10.7554/eLife.45133>
- Barry, S. (2018) Chemists, It is time to embrace preprints. *Chemistry of Materials*, 30, 2859-2859. <https://doi.org/10.1021/acs.chemmater.8b01360>
- Inglis, J.R. & Sever, R. (2016) bioRxiv: a progress report. ASAPBio blog. Accessed June 2019. <https://asapbio.org/biorxiv>
- Monthly Statistics for December 2018. PrePubMed. Accessed June 2019. http://www.prepubmed.org/monthly_stats/
- Narock, T. & Goldstein, E.B. (2019) Quantifying the Growth of Preprint Services Hosted by the Center for Open Science. *Publications*, 7, 44. <https://doi.org/10.3390/publications7020044>
- Polka, J. (2017) Preprints as a complement to the journal system in biology. *Information Services & Use*, 37, 277-280. <https://doi.org/10.3233/ISU-170849>
- Rawlinson, C. & Bloom, T. (2019) New preprint server for medical research. *BMJ*, 365, l2301. <https://doi.org/10.1136/bmj.l2301>
- Submission Guide. bioRxiv. Accessed June 2019. <https://www.biorxiv.org/submit-a-manuscript>
- Teixeira da Silva, J.A. & Dobránszki, J. (2019) Preprint policies among 14 academic publishers. *The Journal of Academic Librarianship*, 45, 162-170. <https://doi.org/10.1016/j.acalib.2019.02.009>
- policies among 14 academic publishers. *The Journal of Academic Librarianship*, 45, 162-170. <https://doi.org/10.1016/j.acalib.2019.02.009>

About the Authors

Rachel Herbert

Rachel Herbert is a Senior Research Evaluation Manager at Elsevier. She has worked in scholarly publishing for over 10 years and has an active interest in the evaluation of research through the lens of gender. Her most recent major project was Elsevier's Research Futures report, which created scenarios of the future of research and research culture over the coming decade.

<https://orcid.org/0000-0002-4088-1223>

Dr. Kate Gasson

Dr. Kate Gasson is a Senior Research Evaluation Manager at Elsevier. She has a Master's degree in Earth Sciences from Oxford University and a PhD in isotope geochemistry from the University of Bristol. Kate left academia in 2015 to pursue a career in publishing, where her work now focuses on developing analytical approaches to derive insights about the world of research using bibliometric and scientometric tools.

<https://orcid.org/0000-0001-5263-146X>

Alex Ponsford

Alex Ponsford is Research Evaluation Manager at Elsevier. He has a BA and MA in History coupled with a background in media analysis. Alex's current interests include the definition and potential measurement of the societal impact of research.

<https://orcid.org/0000-0003-2560-572X>

How to cite:

Herbert, R., Gasson, K. & Ponsford, A. (2019) The Need for Speed, How quickly do preprints become published articles? *ICSR Perspectives*



About the International Center for the Study of Research

The ICSR is tasked with reviewing and advancing the evaluation of research across all fields of knowledge production. Working closely with the research community, the Center draws on interconnected disciplines of research evaluation, bibliometrics and scientometrics, science of science, science and technology studies, and the science of team science to advise, (co)develop and share knowledge within, across and beyond these areas.

Robust, carefully used indicators can help students, faculty, researchers, research administrators and policy makers make the most of the resources at their disposal to achieve their research aims. Smart indicators also help accurately showcase research impact to the global community. On this basis, the Center will identify, review, develop and foster the use of rich and precise qualitative and quantitative indicators of research inputs, activities, outputs and outcomes.

The ICSR works in partnership with a geographically diverse advisory board comprised of experts in research, research evaluation, policy and research management.

Learn more and sign up for email alerts:

www.elsevier.com/icsr

 [@IntCtrStudyRes](https://twitter.com/IntCtrStudyRes)



Scopus is a service mark of Elsevier Inc.
Copyright © 2019 Elsevier B.V. September 2019