Biesenbender, Kristin; Toepfer, Ralf; Peters, Isabella

Experience with posting preprints during the COVID-19 pandemic

ISSI 2023 Proceedings, v2, pp 59–65
DOI: 10.5281/zenodo.8350528 https://cns-iu.github.io/workshops/2023-07-02_issi/

Experience with posting preprints during the COVID-19 pandemic

Kristin Biesenbender[1], Ralf Toepfer[2], Isabella Peters[3]

[1]*k.biesenbender@zbw-online.eu,* [2]*r.toepfer@zbw-online.eu*
ZBW – Leibniz Information Centre for Economics, Neuer Jungfernstieg 21, D-20354 Hamburg, (Germany)

[3]*ipe@informatik.uni-kiel.de*
Kiel University, Christian-Albrechts-Platz 3, D-24105 Kiel, (Germany)

Abstract
In the COVID-19 pandemic, it became much more critical for many life science researchers to rapidly disseminate preliminary versions of their results from one day to the next. This kind of "save the world" feeling prompted a search for ways to circumvent the lengthy publication procedures for journal articles, and many researchers decided to post preprints as upstream publication opportunities. This was a new phenomenon in the life sciences, so it is interesting to examine this development and investigate whether this is a temporary phenomenon or the beginning of a cultural shift in publishing behavior in the life sciences. Preliminary findings from a survey of researchers who have posted COVID-19-related preprints suggest that the large prevalence of preprints in the pandemic could lead to a lasting change in publishing behavior.

Introduction
At the onset of the COVID-19 pandemic, there was a rapid increase in the posting of preprints in the life sciences (Fraser, Brierley, et al., 2021; Puebla et al., 2022). In these communities in particular, there were many pandemic-related topics that scientists wanted to disseminate immediately to swiftly expand knowledge about the impact of the virus or vaccine development (Waltman et al., 2021). Rapid dissemination of scientific articles has clear advantages, especially in times of a pandemic - given the sometimes lengthy publication timelines (Fraser et al., 2020). The publication of preprints on repositories has been recognized as an upstream step before official publication in a journal to increase the immediate visibility of research results. This was made possible by the establishment of preprint repositories such as bioRxiv, medRxiv, and Research Square in 2013, which were initially slow to be adopted by the scientific community. Then, at the onset of the pandemic, there was a rapid increase in the number of preprints on the repositories in April 2020 (Fraser, Brierley, et al., 2021). This is notable because there has been no tradition of posting preprints or working papers in the life sciences, as there is in other disciplines such as physics or economics (Chiarelli et al., 2019).

Objectives
Against this background, we have investigated the experiences of scientists who posted preprints during the COVID-19 pandemic. In addition, we sought to gain insights into whether this proliferation of preprints in the life sciences - after having a pandemic bringing an external shock to the scholarly publication system - has marked the beginning of a cultural shift regarding posting of preprints or whether it is a temporary phenomenon. This leads us to the research question: What is the impact of COVID-19 preprint practices on scholarly publishing behavior? In particular, we are interested in whether the COVID-19 effect on the number of preprints is sustainable and whether authors intend to continue to post preprints after the pandemic. It is also interesting to see whether preprint publication behavior is comparable to that happening during previous health crises. In addition, we wanted to find out whether such "external shocks" are more effective drivers of open access than, for example, mandates. Also we investigated what motivated authors to post preprints during the COVID-19 pandemic.

This research in progress-paper presents preliminary results from an online quantitative survey conducted in September 2022 to learn about researchers' experiences with preprint posting during the COVID-19 pandemic. This is research in progress, because only descriptive statistics can be presented at this point of the analysis. Further bibliometric analysis based on the survey supplemented with preprint, citation, and altmetrics analysis will follow to generate further insights into the impact of preprints in the COVID-19 pandemic.

Methods and data

To gain initial insights to answer our research question, we conducted a quantitative survey of authors of COVID-19 related preprints in September 2022. The 25 survey questions were grouped into five parts: a) experiences with posting preprints during the COVID-19 pandemic, b) motivations for posting preprints, c) concerns about posting preprints, d) role of research institutions or research funders and e) future development. The survey contained single and multiple choice questions, partially supplemented by a free text-option.

To automatically extract preprint data from open data sources (e.g., Crossref and DataCite) a preprint tracker was developed based on keyword queries (Fraser & Kramer, 2020). We thus extracted COVID-19 related preprints from bioRxiv, medRxiv, and Research Square. For bioRxiv and medRxiv, we extracted 24,436 COVID-19 related preprints for which we obtained 51,335 email addresses by scraping from repositories websites. After cleaning, 22,219 email addresses from bioRxiv and medRxiv remained. Two thousand additional email addresses from among a total of 11,194 randomly selected COVID-19-related preprints on Research Square were manually collected, for which we could not find an automated solution. Finally, we sent the survey to 24,219 authors of COVID-19 preprints. In the end, we received 1,131 completed responses (response rate: 4.9%), which formed the basis for our analysis.

The majority of respondents came from the United States (25%), followed by the United Kingdom (10%) and Germany (7%). India, Canada, Spain, Italy and France were next in line. Using the Frascati Classification on Science and Technology (OECD, 2002) for self-assignment of disciplines, 666 respondents stem from the medical sciences, followed by 202 from the natural sciences, 101 from the social sciences and 81 from engineering and technology. The response to the question “How many years have you been doing research?” show that there are many respondents with quite long research experience. 407 respondents reported doing research for five to 14 years (36%), 306 for 15 to 24 years (27%), and 298 for more than 24 years (26%). 776 respondents were from universities, 240 from non-university research institutions.

Results

Experience with posting preprints during the COVID-19 pandemic

To learn whether the researchers had experience posting preprints prior to the COVID-19 pandemic, we asked them, “When was the first time you posted a paper on a preprint server/repository?” The vast majority (69%) reported posting a preprint for the first time during the pandemic (i.e., in the years 2020, 2021 and 2022; Table 1).

Table 1. First time of posting a paper on a preprint server/repository.

	FirstTime_count	*FirstTime_percent*
Before 2020	347	30.7%
During 2020	480	42.4%
During 2021	252	22.3%
This year	50	4.4%
N/A	2	0.2%
Sum	1131	100%

We also asked, "How many preprints have you posted since the pandemic?" 81% said that they had published one to five preprints since the pandemic outbreak, and about 6% even said that they had published 10 or more preprints. In one of the survey questions, we asked for comparison with previous pandemics. Only 76 respondents answered, while 1055 respondents skipped the question. 1.8% of respondents have posted preprints on the ZIKA virus, 1.3% on Ebola virus, and 0.7% on swine flu. This hints towards preprints having not played any role in previous health crises. In addition, we wanted to learn more about the researchers' experiences with posting preprints during the COVID-19 pandemic, particularly in terms of visibility and citations received from the scientific community, attention from the broader public and media, and social media (Figure 1). About 65% (strongly) agreed that their preprints were particularly visible to the scientific community during the pandemic. 47% (strongly) agreed that their preprints received a lot of attention from a broad audience and/or the media, and that their preprints were shared more frequently via social media. Finally, 43% (strongly) agreed that their preprints were cited more frequently during the pandemic.

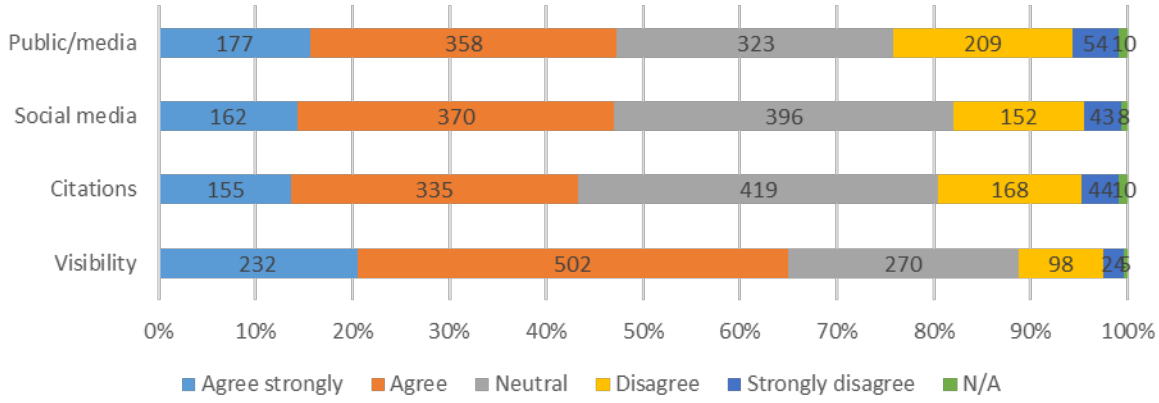

Figure 1. Experience with posting preprints during the COVID-19 pandemic in terms of visibility in the scientific community, the broader public/media, and social media.

Furthermore, we were interested in other potential benefits that researchers experienced from posting preprints during the pandemic, namely availability, public discussions, collaboration, journal invitations, and quality of preprints (Figure 2). Most researchers agreed that posting preprints during the pandemic was essential to make their findings directly available to other researchers (84% agreed/strongly agreed) as well as to society, which 67% (strongly) agreed to be essential. Nearly 50% (strongly) agreed that posting preprints led to invitations by journal editors to publish their preprint in an academic journal, and 45% indicated that it encouraged collaborations with other researchers. Diverging experiences were found for whether posting

preprints during the pandemic helped to ensure the quality of their publication: 30% (strongly) agreed and 30% (strongly) disagreed, respectively.

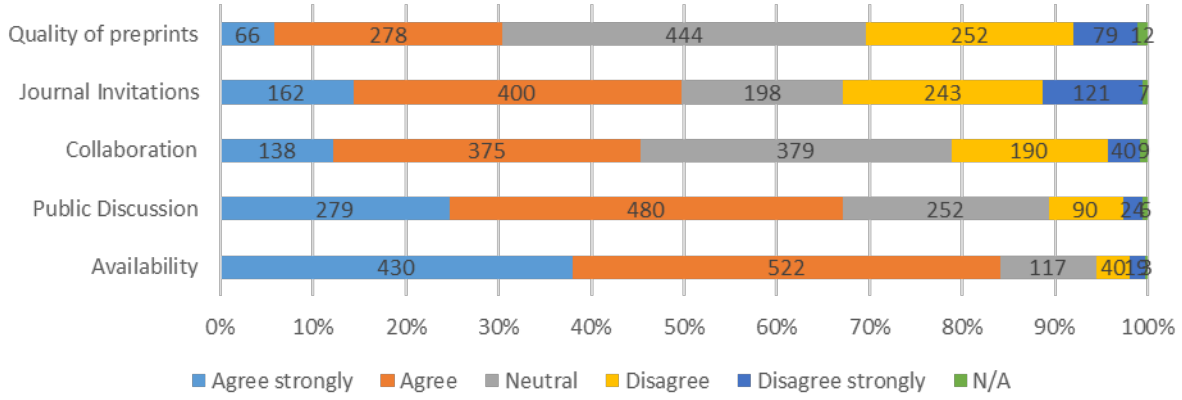

Figure 2. Experience with posting preprints during the COVID-19 pandemic in terms of availability, public discussions, collaboration, journal invitations, and quality of preprints.

In addition, we were interested in whether there is any content quality control of preprints and where this takes place. About 57% of the respondents answered that there is always some kind of content quality control before their preprints are posted on a preprint server/repository. If we add the 13% of respondents who answered that content quality control sometimes occurs, we can say that about 70% of preprints go through at least some kind of quality control. 56% reported that the quality control takes place in their research group, 43% answered that it happens on the preprint server.

Motivations for posting a preprint

By far the most important motivation for researchers to post a preprint is to increase awareness of their research (76%). The second most important motivation for posting a preprint was to "stake a claim on my findings" (52%). Responses to "receive early feedback" are fairly balanced. For some it is a motivation, for others it is not. Promoting research collaborations and publishing more detailed versions of a paper were not strong motivations.

Concerns regarding posting a preprint

Lack of quality assurance (45%) and the risk of the media reporting incorrect results (42%) are the most commonly reported concerns about posting a preprint on a preprint server. In addition, only 15% of respondents stated that: "A manuscript should only be published in a scientific journal if its content has not yet been posted as a preprint". Conversely, this means that 85% of the respondents answered that the "Ingelfinger rule" no longer applies. Furthermore, only about 10% of respondents answered that comments and feedback are generally not helpful.

Role of research institutions and research funders

To examine the role of research institutions and funders during the pandemic, we asked whether it was necessary for researchers to comply with an institution's or funder's open access/preprint policy regarding preprint publication. 77% of respondents indicated that it was not necessary. And 81% pointed out that this was also true for funding agencies. Very few researchers reported that research funders or their own institution facilitated preprint publication.

Future development and sustainability

To answer our research question above, we are interested in whether the increase in posting preprints is sustainable and how it will develop in the future. To investigate this, we asked researchers if they intend to post their scholarly work on preprint servers/repositories in the future. 23% responded that they will post all of their future work, and 56% will post some of their future work. In addition, we asked about the drivers that researchers believe are causing the increased posting of preprints in their field (Figure 3). 78% (strongly) agreed with the statement that external shocks are powerful drivers. In comparison, only 46% (strongly) agreed with this statement regarding mandates or policies and even less regarding the disciplinary culture. More generally, we were interested in what the researchers thought about a cultural shift in publishing behavior in their field, namely whether posting preprints on a server will become common practice in their field, to which 67% (strongly) agreed. More specifically, we asked whether overlay journals could be a solution to combine the advantage of early access with improved quality control. 61% of respondents (strongly) agreed.

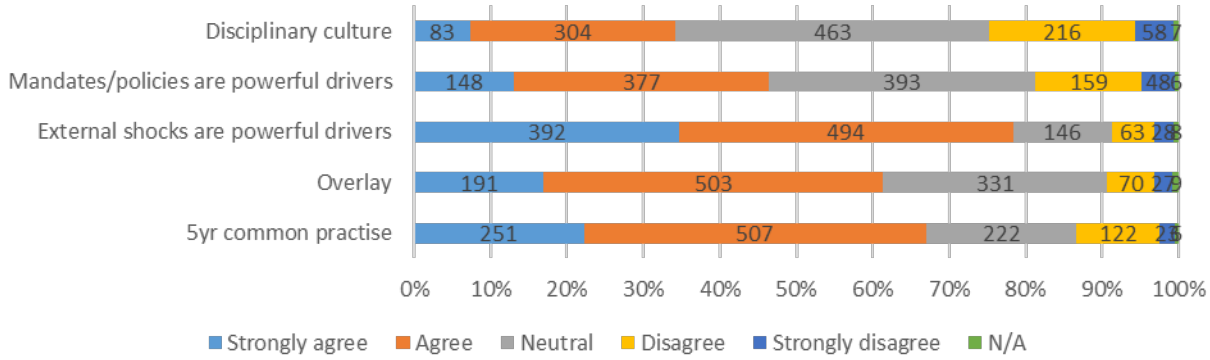

Figure 3. Sustainability of posting preprints in the future.

Discussion

We aim at understanding the impact of COVID-19 on preprint practices and on scholarly publishing behavior. In this regard, we found that there are early indications that the COVID-19 pandemic has triggered a significant increase in the number of preprints, which, while not remaining at the same level, seems to be translating into a sustained change in publication behavior in the life sciences. From our survey, the vast majority (69%) posted a preprint for the first time during the pandemic. This is consistent with other studies that have shown an increase in preprints at the beginning of the pandemic (Fraser, 2021; Waltman et al., 2021). The survey results also clearly show that preprint posting behavior is not comparable to that from previous health crises where only few respondents published preprints to disseminate results quickly.

Furthermore, with 79% of respondents indicating that they will post at least some of their future work, which was also found by others (e.g. Waltman et al., 2021). Thus, we have preliminary evidence that authors who posted preprints during the COVID-19 pandemic will continue to do so. However, respondents reported good experiences with posting preprints, e.g. many researchers experienced that their preprints were particularly visible during the pandemic. They pointed out the importance of making their findings directly available to other researchers, which has been shown in other studies (Fraser et al., 2022; Waltman et al., 2021).

We can conclude that the "external shock" related to COVID-19 has been the most effective driver of open access from the perspective of the responding researchers. It was stronger than mandates, as respondents predominately indicated that they did not have to comply with institutions' or funders' policies during the pandemic. In addition, many confirmed that pre-publication as a preprint does not stand in the way of publication in a journal. In terms of future developments, it is also interesting to note that almost 50% of respondents (strongly) agreed that posting preprints during the pandemic led to invitations from journal editors to publish their preprint in a scientific journal. This could be an indication of the greater emergence of overlay journals (Rousi & Laakso, 2022). But, it could also be that there are more (sometimes of lower quality or predatory) journals that offer to publish preprints without peer review in exchange for an article processing charge. Further research is required here.

In addition, we found differing opinions about whether publishing preprints improves the quality of research. As others have found, there are concerns about quality control as well as the presentation of scientific information to the public that has not undergone peer review (Fraser et al., 2022; Penfold & Polka, 2020). However, 70% of respondents pointed out that preprints are subject to at least some content quality control carried out by the scientific community. Considering that a lack of quality control is one of the main reasons for concerns about preprint publishing, we find this result quite remarkable. It might be interesting for further research to learn more about how exactly this quality control occurs.

Conclusion

The COVID-19 pandemic was a powerful driver of preprint publication in the life sciences. Although external shocks have been more effective in driving change, the survey respondents believed that mandates and policies could support this development in the future as can different publication models such as overlay journals. Further research would need to investigate whether there are correlations between responses and demographics, especially with countries, affiliations, and years of research activity. In addition, free text responses from the partially supplemented open categories in the questions should be analyzed for further insights, for example, regarding motivations and concerns for posting preprints. And finally, an output analysis of preprints during the pandemic will be conducted to examine the sustainability of the COVID-19-shock. Complemented by a citation and altmetrics analysis, we aim to explore the impact of the COVID-19 pandemic on citation and online dissemination practices for preprints.

Acknowledgments

This work is funded by the German Federal Ministry of Education and Research (BMBF) project Open Access Effects – The Influence of Structural and Author-specific Factors on the Impact of OA (OASE), grant number 16PU17005A/16PU17005B.

Data availability

Questionnaire and response data used for the preparation, analysis and visualization of response data are available on SowiDataNet/datorium: https://doi.org/10.7802/2552. Note that raw free-text responses and email addresses of survey respondents were removed to preserve participant anonymity.

References

Chiarelli, A., Johnson, R., Pinfield, S., & Richens, E. (2019). *Accelerating scholarly communication: The transformative role of preprints*.

Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálfy, M., Nanni, F., & Coates, J. A. (2021). The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLOS Biology*, *19*(4).

Fraser, N., & Kramer, B. (2020). *Covid19_preprints [Data Set]*. Figshare.

Fraser, N., Mayr, P., & Peters, I. (2022). Motivations, concerns and selection biases when posting preprints: A survey of bioRxiv authors. *PLOS ONE*, *17*(11).

Fraser, N., Momeni, F., Mayr, P., & Peters, I. (2020). The relationship between bioRxiv preprints, citations and altmetrics. *Quantitative Science Studies*, *1*(2), 618–638.

OECD. (2002). *Frascati Manual 2002: Proposed Standard Practice for Surveys on Research and Experimental Development*. Organisation for Economic Co-operation and Development.

Penfold, N. C., & Polka, J. K. (2020). Technical and social issues influencing the adoption of preprints in the life sciences. *PLOS Genetics*, *16*(4).

Puebla, I., Polka, J., & Rieger, O. Y. (2022). *Preprints: Their Evolving Role in Science Communication*. Against the Grain (Media), LLC.

Rousi, A. M., & Laakso, M. (2022). *Overlay journals: A study of the current landscape*. arXiv.

Waltman, L., Pinfield, S., Rzayeva, N., Oliveira Henriques, S., Fang, Z., Brumberg, J., Greaves, S., Hurst, P., Collings, A., Heinrichs, A., Lindsay, N., MacCallum, C. J., Morgan, D., Sansone, S.-A., & Swaminathan, S. (2021). *Scholarly communication in times of crisis: The response of the scholarly communication system to the COVID-19 pandemic*. Research on Research Institute.