Is the acceptance time shorter for submission with preprints?

Dan Tian, Xin Liu, and Jiang Li

DG20140026@ smail.nju.edu.cn; liuxin@smail.nju.edu.cn; lijiang@nju.edu.cn
0000-0003-1980-8467; 0000-0002-0399-6611; 0000-0001-5769-8647
School of Information Management, Nanjing University, China

Abstract: Preprints are becoming increasingly vital in scientific communication. This study aims to investigate whether preprints offer any advantages regarding acceptance time for researchers, based on the characteristics of early accessibility and peer feedback through preprints. We employ regression analysis to compare the acceptance time difference between manuscripts with preprints and those without preprints submitted to the same journal in the same year and month. The findings reveal that manuscripts that had preprints released before submission experienced significantly shorter acceptance times than those submitted to the same journal in the same month and year without preprints. However, this advantage did not persist when preprints were posted after submission.

1. Introduction

The publication process in academia is progressively decelerating (Ellison, 2002). Although digital publishing may offer advantages in terms of expediting the publication timeline, it does not necessarily shorten the time invested in peer review (Powell, 2016). The anticipated duration for manuscript review by the author is eight weeks, yet the reality is that it typically takes fourteen weeks (Nguyen et al., 2015). Journal editors are faced with handling a greater volume of submissions due to an increase in submissions, while reviewers now require more data, revisions, and new experiments than before.

To avoid time delays related to peer review and publication, numerous researchers in diverse fields disseminate their latest research findings through preprints. Preprints represent manuscripts that have not yet undergone formal peer review but are available to the public online via preprint servers (Berg et al., 2016). Preprints enable research findings to receive earlier and faster attention from peers, increasing the visibility of research outcomes (Fraser et al., 2020; Serghiou & Ioannidis, 2018). Furthermore, preprints have the potential to enhance research quality (Sarabipour et al., 2019). Preprint platforms offer researchers an opportunity to receive feedback from their peers before undergoing formal peer review, enabling earlier identification of potential inaccuracies and flaws (Anderson et al., 2014). By engaging in feedback and discussions with other researchers, authors can refine their research methods and strengthen their conclusions, ultimately improving the quality of their work.

Based on the characteristics of early accessibility and peer feedback via preprints, we hypothesize,
H1: *The acceptance time is shorter for submissions with preprints.*
On the one hand, the journal editor or reviewer may have read the study on the preprint platform before receiving the manuscript, knowing the study's strengths and weaknesses and making a swift decision without having to review it from scratch. On the other hand, peer feedback on the preprint platform can enhance manuscript quality, resulting in fewer revisions being necessary during the review process.

2. Literature review

The arXiv was established in the early 1990s as a preprint platform for physics and mathematics, and has since expanded to include other fields, such as computer science, biology, and statistics (Berg, 2017). In 2013, bioRxiv was launched as a preprint server for biology, and

has become a popular platform for sharing research in life sciences. Other preprint servers have also emerged, including medRxiv for medical research. The COVID-19 pandemic has increased the attention given to preprints, as many scientists turned to preprint servers to rapidly share their research in response to the urgent need for information (Fraser et al., 2021).

Numerous studies have examined factors related to acceptance time. There are significant variations in acceptance time among different journals (Runde, 2021; Sebo, 2023). Previous research has shown that submission rates may vary seasonally, with the top psychology journal *Psychological Science* receiving the highest number of submissions during the summer (Shalvi et al., 2010). Editors and reviewers may face more submissions during certain months (Ausloos et al., 2019; Schreiber, 2012), which could affect the speed of manuscript processing and thus impact acceptance time. Additionally, the number of authors positively correlates with acceptance time, and papers from authors in high-income countries have an advantage regarding acceptance time (Taşkın et al., 2022).

The most relevant study to this research was conducted by Tsunoda et al. (2020). They compared the acceptance time difference between papers first posted on bioRxiv before being submitted to *PLoS One* and those first submitted to *PLoS One* and later posted as preprints. Their study only used papers with preprints published in *PLoS One*, and the authors acknowledged the need to validate their findings in a broader range of journals. Their another study compared the acceptance time of papers with and without preprints in 119 journals (Tsunoda et al., 2022). They found that papers with preprints published in 29 journals had significantly shorter acceptance time than those without preprints. However, they did not account for other factors that could have influenced acceptance time, such as submission dates, the number of authors, and the authors' country, all of which could have affected the results. In contrast to prior studies, our research has significant updates to both data and methodology. We analyse preprints posted on arXiv and medRxiv in addition to bioRxiv and conducted a regression analysis to control for a range of factors related to acceptance time based on matched samples of papers with and without preprints submitted to the same journal in the same year and month.

3. Methods

To mitigate the influence of journal and submission time on paper acceptance, we conducted a matching procedure between papers with and without preprints submitted to the same journal within the same year and month, exclusively selecting journal articles. Preprint data was sourced from three platforms, arXiv, bioRxiv, and medRxiv, and publication timeline information was gathered from PubMed. To our knowledge, PubMed is the only bibliographic database that provides publication timeline information for papers.

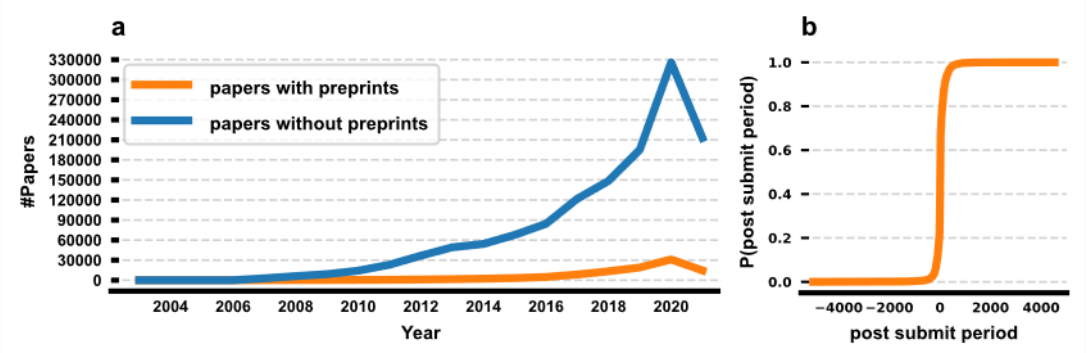
The process of sample matching comprised four primary stages. Firstly, we retrieved preprint metadata via APIs from three platforms, collecting metadata up to 2022, including titles, authors, preprint unique identifiers, and DOIs after publication. Secondly, we checked whether PubMed had indexed published preprints, matching the DOI of the preprint after publication with the bibliographic data available on PubMed. Thirdly, we obtained preprint samples with submission and acceptance time data and calculated the acceptance time. Lastly, we matched papers submitted without preprints to the same journal within the year and month. The sample included 101,676 papers with preprints and 1,352,691 papers without preprints, representing 2,795 journals. Table 1 illustrates the sample size at each stage of the screening process.

Table 1. The number of preprints corresponding to the three platforms

	arXiv	**bioRxiv**	**medRxiv**	**Total**
#preprints	2,166,248	177,913	37,514	2,381,675
#preprints assigned post-publication DOI	1,543,093	97,704	20,310	1,661,107
#preprints index in PubMed	94,613	75,203	11,769	181,585
#preprints with received and accepted date	37,591	61,439	9,020	108,050
#preprints matched papers without preprints	35,265	58,308	8,103	101,676

Figure 1a displays the number of papers submitted annually in the sample. The number of papers submitted without preprints each year significantly surpasses those submitted with preprints. Figure 1b illustrates the duration between the preprint's posting on the platform and its submission to the journal. Two types of preprints can be distinguished based on their order of posting on the platform and submitting to a journal. Manuscripts first posted on a preprint server and subsequently submitted to a journal (post submit period > 0) account for 55%, while manuscripts first submitted to a journal and then uploaded to a preprint server (post submit period $\leqslant$ 0) account for 45%. In particular, 55,518 papers were posted on the preprint platform before submission and matched 1,109,439 non-preprint papers published in the same journal in the same year and month. In contrast, 46,158 papers were released as preprints after submission and matched 1,011,762 non-preprint papers published in the same journal in the same year and month.

Figure1. The number of papers submitted annually and the distribution of post submission period



We use the model shown in Model (1) to evaluate whether the papers with preprint have an advantage in acceptance time.

$$acceptance\_time = \beta_1 has\_prerint + \varphi X + \delta + \varepsilon \quad (1)$$

where *acceptance_time* represents the time a paper takes to be accepted from submission. *has_preprint* indicates whether a paper has a preprint (yes = 1, no = 0). X represents a series of variables related to the acceptance time, including the number of authors, references, and the corresponding author's country. δ represents the fixed effects of the journal and submission date for a paper.

Considering that different preprints may have varying probabilities of being read by peers, we grouped the samples for regression analysis. Firstly, the quality of preprints can influence the likelihood of being read by peers, with higher-quality preprints having a greater chance of being seen. Therefore, we grouped the samples based on the 2021 Journal Impact Factor Quartile in which the preprint was published. Secondly, the duration of the manuscript on the preprint platform can impact the probability of being read. Typically, the longer a preprint remains on a preprint platform, the higher the likelihood of it being read by peers. As a result, we divided the

sample into four groups based on the quartiles of the period between the preprint's release date and the journal submission date. Lastly, the prominence, maturity, and research field can vary across different preprint platforms, resulting in differing chances of preprints posted on different platforms being read by peers. Therefore, we grouped and regressed the samples based on the preprint publication platform.

4. Results

Table 2 presents the descriptive statistics of the variables utilized in this research. The dependent variable is *acceptance_time*, which ranges from 0 to 4,043, with a mean of 119.96 and a standard deviation of 93.02. The variable *has_preprint* has a mean of 0.07, indicating that 7% of the analyzed papers have preprints. Furthermore, we presented the statistics of control variables associated with acceptance time.

Table 2. Descriptive statistics of the variables

	acceptance_time	has_preprint	#Author	#Reference
Obs.	1,454,367	1,454,367	1,454,367	1,454,367
Mean	119.96	0.07	7.16	46.53
Std. Dev.	93.02	0.25	22.11	36.76
Min	0	0	1	0
Max	4,043	1	2,960	2,352

The results in columns (1)-(2) of Table 3 are based on the samples of papers released as preprints before submission to journals and their corresponding counterparts without preprints. The results in columns (3)-(4) of Table 3 are based on the samples of papers posted on preprint platforms after submission to a journal and their matched papers without preprint.

We included fixed effects for journal and submission time in all models. Column (1) does not add any control variables, and the coefficient for *has_preprint* is negative and statistically significant. In Column (2), we added all control variables, and the coefficient for *has_preprint* is -7.277 ($p<.001$), suggesting that papers with preprints have a 7.277-day shorter acceptance time than those without preprints. However, the coefficients for *has_preprint* in Columns (3) and (4) are positive and statistically significant, indicating that papers with preprints do not have an advantage in acceptance time compared to those without preprints.

Table 3. The difference in acceptance time between papers with preprints and papers without preprints

	Preprints post before submission		Preprints post after submission	
	(1)	(2)	(3)	(4)
has_preprint	-7.796***	-7.277***	5.387***	5.829***
	(0.352)	(0.354)	(0.481)	(0.482)
#Author		0.047***		0.109***
		(0.007)		(0.013)
#Reference		0.028***		0.039***
		(0.002)		(0.002)
Corresponding author's country		√		√
Journal-Submitted_year_month fixed effects	Yes	Yes	Yes	Yes
Observations	1,164,956	1,164,947	1,057,918	1,057,906

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In Table 4, by comparing the acceptance time of manuscripts with preprints and those without preprints submitted in the same year and month in the same journal, we further analyze the advantages of preprints regarding acceptance time across different journal quartiles. Table 4 shows that the coefficients of *has_preprint* in all models are statistically significant and negative, suggesting that papers with preprints published across various JIF Quartile experience significantly shorter acceptance time than those without preprints.

Table 4. Regression results for samples of different JIF Quartile

	acceptance_time			
	JIF Q1	JIF Q2	JIF Q3	JIF Q4
has_preprint	-5.407***	-8.915***	-4.183***	-11.584***
	(0.452)	(0.431)	(0.920)	(2.352)
#Author	0.043***	0.024***	0.368***	0.394**
	(0.009)	(0.005)	(0.060)	(0.160)
#Reference	0.056***	0.004*	0.018***	0.109***
	(0.002)	(0.002)	(0.007)	(0.027)
Corresponding author's country	√	√	√	√
Journal-Submitted_year_month fixed effects	Yes	Yes	Yes	Yes
Observations	577,357	837,023	124,228	12,961

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

According to Table 5, when the manuscript is posted on a preprint platform within 13 days before journal submission, the acceptance time for papers with a preprint is 4.331 days less than papers without a preprint. Similarly, when the periods are 13~50, 50~145, and 145~4,534, the average time for accepting papers with preprints is 3.363 days, 9.315 days, and 10.084 days shorter than that of papers without preprints, respectively.

Table 5. Regression results for samples of different duration in preprint platforms

	acceptance_time			
	1≤period≤13	13<period≤50	50<period≤145	145<period<4,534
has_preprint	-4.331***	-3.363***	-9.315***	-10.084***
	(0.666)	(0.679)	(0.674)	(0.693)
#Author	0.036***	0.199***	0.299***	0.354***
	(0.006)	(0.040)	(0.050)	(0.052)
#Reference	0.035***	0.037***	0.032***	0.033***
	(0.003)	(0.003)	(0.003)	(0.003)
Corresponding author's country	√	√	√	√
Journal-Submitted_year_month fixed effects	Yes	Yes	Yes	Yes
Observations	672,827	685,603	724,808	735,536

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6 displays the time difference for acceptance between preprints posted on arXiv, bioRxiv, and medRxiv, and papers without preprints. It shows that the coefficients for has preprint are

negative and statistically significant, indicating a shorter acceptance time for papers with preprints posted on any of the three preprint platforms compared to those without preprints.

Table 6. Regression results for samples of different preprint platforms

	acceptance_time		
	arXiv	bioRxiv	medRxiv
has_preprint	-13.935***	-4.036***	-9.546***
	(0.682)	(0.450)	(0.885)
#Author	0.018***	0.536***	0.366***
	(0.005)	(0.026)	(0.027)
#Reference	0.059***	0.026***	0.022***
	(0.005)	(0.002)	(0.003)
Corresponding author's country	√	√	√
Journal-Submitted_year_month fixed effects	Yes	Yes	Yes
Observations	593,155	840,781	243,566

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5. Discussion

This study investigates whether preprints confer an advantage in acceptance time by comparing the differences between papers with and without preprints. The results indicate that papers with preprints released before submission had significantly shorter acceptance time than those without preprints submitted to the same journal in the same month and year. However, posting preprints after submission did not confer this advantage.

The study finds that papers with preprints are accepted faster than those without preprints across journals of different tiers, and there is no evidence to suggest that preprints confer different advantages regarding acceptance time for different journal tiers. When a manuscript has been available on preprint platforms for less than 50 days, papers with preprints are accepted 3-4 days faster than those without preprints. For manuscripts available for over 50 days, papers with preprints are accepted 9-10 days earlier than those without. Preprints released on all three platforms confer advantages regarding acceptance time, with the most significant reduction in acceptance time observed for preprints posted on arXiv.

The advantage in acceptance time for papers with preprints is significant, with papers having preprints accepted 7.277 days faster than those without preprints submitted to the same journal in the same month and year. This finding has important implications for researchers and scientific development. As per Merton's norms of science, scientific discoveries are not the private property of scientists, and recognition and respect are the only rights they have over their discoveries (Merton, 1995). In academia, there are still many countries in which journal publications are valued more than preprints in research evaluation. Therefore, faster acceptance signifies an earlier recognition from peers.

This study does not establish a cause-and-effect relationship between preprints and acceptance time, as there may be other competing explanations. For instance, it is possible that researchers selectively post their work of high quality on preprint platforms. If manuscripts posted on preprint platforms are inherently of higher quality, the advantage of preprints in terms of acceptance time may not be due to increased visibility or improved research quality but rather to the higher quality of the studies themselves. Although the available evidence suggests that the authors did not selectively publish high-quality research on the preprint platform (Fraser et

al., 2022), further research is necessary to rule out this explanation. Additionally, submitting a paper for publication typically involves three stages: the editor's decision time, the reviewer's review time, and the author's response time (Huisman & Smits, 2017), with the time spent on each stage potentially impacting the acceptance time. Due to data limitations, we did not investigate at which stage the preprint has advantages in the publishing process, and this is worth exploring further.

Open science practices

The metadata of PubMed, arXiv, bioRxiv, and medRxiv are openly available. We access and download PubMed baseline data through https://pubmed.ncbi.nlm.nih.gov/download/#annual-baseline. We utilize public APIs to access the metadata of arXiv, bioRxiv, and medRxiv, respectively. We use Python to process and analyze data. Python codes are available from the corresponding author upon request.

Acknowledgments

We thank Mr. Zhenyue Zhao for his valuable suggestions on this article.

Author contributions

Dan Tian: Conceptualization, Methodology, Software, Data curation, Writing - original draft. Xin Liu: Data curation, Methodology. Jiang Li: Supervision, Conceptualization, Methodology, Writing - original draft, Writing - review & editing.

Competing interests

The authors declare no competing interests.

Funding information

This research has received external funding from the Postgraduate Research & Practice Innovation Program of Jiangsu Province.

References

Anderson, N., Potocnik, K., & Zhou, J. (2014). Innovation and Creativity in Organizations: A State-of-the-Science Review, Prospective Commentary, and Guiding Framework. *Journal of Management*, *40*(5), 1297-1333. https://doi.org/10.1177/0149206314527128

Ausloos, M., Nedič, O., & Dekanski, A. (2019). Correlations between submission and acceptance of papers in peer review journals. *Scientometrics*, *119*(1), 279-302. https://doi.org/10.1007/s11192-019-03026-x

Berg, J. M. (2017). Preprint ecosystems. *Science*, *357*(6358), 1331-1331. https://doi.org/doi:10.1126/science.aaq0167

Berg, J. M., Bhalla, N., Bourne, P. E., Chalfie, M., Drubin, D. G., Fraser, J. S., . . . Wolberger, C. (2016). Preprints for the life sciences. *Science*, *352*(6288), 899-901. https://doi.org/10.1126/science.aaf9133

Ellison, G. (2002). The slowdown of the economics publishing process. *Journal of Political Economy*, *110*(5), 947-993. https://doi.org/10.1086/341868

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), e3000959. https://doi.org/10.1371/journal.pbio.3000959

Fraser, N., Mayr, P., & Peters, I. (2022). Motivations, concerns and selection biases when posting preprints: A survey of bioRxiv authors. *PloS One*, *17*(11), e0274441. https://doi.org/10.1371/journal.pone.0274441

Fraser, N., Momeni, F., Mayr, P., & Peters, I. (2020). The relationship between bioRxiv preprints, citations and altmetrics. *Quantitative Science Studies*, *1*(2), 618-638. https://doi.org/10.1162/qss_a_00043

Huisman, J., & Smits, J. (2017). Duration and quality of the peer review process: the author's perspective. *Scientometrics*, *113*(1), 633-650. https://doi.org/10.1007/s11192-017-2310-5

Merton, R. K. (1995). The Thomas theorem and the Matthew effect. *Social Forces*, *74*(2), 379-422. https://doi.org/10.2307/2580486

Nguyen, V. M., Haddaway, N. R., Gutowsky, L. F., Wilson, A. D., Gallagher, A. J., Donaldson, M. R., . . . Cooke, S. J. (2015). How long is too long in contemporary peer review? Perspectives from authors publishing in conservation biology journals. *PloS One*, *10*(8), e0132557. https://doi.org/10.1371/journal.pone.0132557

Powell, K. (2016). Does it take too long to publish research? *Nature*, *530*(7589), 148-151. https://doi.org/10.1038/530148a

Runde, B. J. (2021). Time to publish? Turnaround times, acceptance rates, and impact factors of journals in fisheries science. *PloS One*, *16*(9), e0257841. https://doi.org/10.1371/journal.pone.0257841

Sarabipour, S., Debat, H. J., Emmott, E., Burgess, S. J., Schwessinger, B., & Hensel, Z. (2019). On the value of preprints: An early career researcher perspective. *PLoS Biology*, *17*(2), e3000151. https://doi.org/10.1371/journal.pbio.3000151

Schreiber, M. (2012). Seasonal bias in editorial decisions for a physics journal: you should write when you like, but submit in July. *Learned Publishing*, *25*(2), 145-151. https://doi.org/https://doi.org/10.1087/20120209

Sebo, P. (2023). Are acceptance and publication times longer in primary health care journals compared to internal medicine journals? A comparative study of 117 high-impact journals. *Scientometrics*, *128*(1), 873-876. https://doi.org/10.1007/s11192-022-04593-2

Serghiou, S., & Ioannidis, J. P. A. (2018). Altmetric Scores, Citations, and Publication of Studies Posted as Preprints. *JAMA*, *319*(4), 402-404. https://doi.org/10.1001/jama.2017.21168

Shalvi, S., Baas, M., Handgraaf, M. J. J., & De Dreu, C. K. W. (2010). Write when hot — submit when not: seasonal bias in peer review or acceptance? *Learned Publishing*, *23*(2), 117-123. https://doi.org/https://doi.org/10.1087/20100206

Taşkın, Z., Taşkın, A., Doğan, G., & Kulczycki, E. (2022). Factors affecting time to publication in information science. *Scientometrics*, *127*(12), 7499-7515. https://doi.org/10.1007/s11192-022-04296-8

Tsunoda, H., Sun, Y., Nishizawa, M., Liu, X., & Amano, K. (2020). The influence of bioRχiv on PLOS ONE's peer-review and acceptance time. *Proceedings of the Association for Information Science and Technology*, *57*(1), e398. https://doi.org/10.1002/pra2.398

Tsunoda, H., Sun, Y., Nishizawa, M., Liu, X., & Amano, K. (2022). How Preprint Affects the Publishing Process: Duration of the Peer Review Process between bioRxiv and Journal Papers. *Proceedings of the Association for Information Science and Technology*, *59*(1), 505-509. https://doi.org/10.1002/pra2.660