Preprint Servers' Policies, Submission Requirements, and Transparency in Reporting and Research Integrity Recommendations

Mario Malicki (✉ mario.malicki@mefst.hr)

Stanford University   https://orcid.org/0000-0003-0698-1930

Ana Jerončić

University of Split School of Medicine   https://orcid.org/0000-0003-1621-1956

Lex Bouter

Amsterdam UMC   https://orcid.org/0000-0002-2659-5482

Gerben ter Riet

Amsterdam University of Applied Sciences   https://orcid.org/0000-0002-2231-7637

John P.A. Ioannidis

Stanford University   https://orcid.org/0000-0003-3118-6859

Steven M. Goodman

Stanford University   https://orcid.org/0000-0002-3872-5723

IJsbrand Jan Aalbersberg

Elsevier   https://orcid.org/0000-0002-0209-4480

Research Article

Abstract

In light of the recent increase in the number of preprints and preprints servers, as well as calls for transparent reporting of research, we conducted a cross-sectional analysis of 57 preprint servers' websites for information on their policies, submission requirements, and transparency in reporting and research integrity recommendations. All servers specified their scope, and a large majority mentioned they apply moderation or screening procedures (n=47, 82%). Out of 18 analysed topics on transparency in reporting and research integrity, the servers addressed a median of 1 (IQR 1 to 3), most commonly data sharing (n=22, 39%), plagiarism (n=15, 26%) and the use of ORCID iD (n=14, 25%). Preprint servers could do more to raise awareness and encourage or require transparent reporting of research and adherence to research integrity standards. In doing so they may improve the quality and trust in the scholarly information exchange.

Introduction

While peer review and scholarly publishing have been a cornerstone of scientific communication for at least a century,[1-3] many criticisms have been raised regarding the lack of transparency or quality assurances they provide,[4-7] and the long delays between submission and publication.[8-10] Experiments with faster dissemination of research began in the 1960s,[11] and in the1990s first preprint servers were created, of which the largest and longest running are arXiv, SSRN and RePec.[12-14] Since 2010, more than 30 new preprint servers have emerged,[15] with additional servers recently announced.[16-18] The number of deposited preprints has grown exponentially in the last five years, especially in the life sciences[19, 20] with numerous journals now supporting posting of preprints and accepting preprints as submissions for journal peer review and publication.[21] Research on preprints is, however, still scarce, with some studies demonstrating an initial increase in citations, tweets or other altmetrics scores for studies that were first shared as preprints.[22-27] In regards to content, two recent studies found very few differences in text and completeness of reporting between preprints and versions of record published in scientific journals.[28, 29]

As transparency and completeness of reporting of studies seems to be associated with the recommendations and policies of the journals they are published in, [30-32] and inspired by our recent exploration of recommendations or requirements listed in journal's *Instructions to Authors*,[33] we aimed to explore what preprint servers recommend or demand in their instructions to authors, especially regarding their policies, submission requirements, and transparency in reporting and research integrity recommendations.

Materials And Methods

We conducted a cross-sectional analysis of preprint servers' webpages accessed between 25 January 2020 and 31 March 2020. We analysed any page resembling *Instructions to Authors* (ItAs) or containing information on the preprint servers and their submission procedures. We did this, by starting from the main page and exploring links presented there, as well as in *About*, *Policy*, or *Frequently Asked Questions* pages. Finally, for each server, Mario Malički created user accounts and went through the preprint submission process (without submitting a preprint) to check if additional information was available in the online submission platforms (except for ChinaXiv which required an email associated with a Chinese institution).

Preprint server selection

We define a preprint as any scholarly manuscript version that has not been peer reviewed or published in a traditional publishing venue (e.g. scientific journal, conference proceeding, book) and that is shared publicly through an online platform.We consider a preprint server, an online platform that enables posting of preprints and provides users (i.e. researchers, stakeholder or the public) the ability to search for, view and download preprints. As there is currently no established list or a database of all existing preprint servers, we compiled our own list using Martyn Rittman's list,[34] ASAPbio list,[35] Wikipedia List of preprint servers,[36] existing StackExchange thread,[37] and personal knowledge. We additionally excluded all servers, platforms or institutional repositories that: 1) limit the ability to deposit preprints only to researchers that received funding from a specific source or had a specific institutional affiliation (e.g. AAS or Wellcome Open Research); 2) actively seek out peer review or function as "in-review" or first look services (e.g. F1000 Research, Cell Press Sneak Peek, Springer Nature In Review); 3) do not have search options for listing of preprints on the platform (e.g. ResearchGate); or 4) no longer accept preprints (e.g. Peer J preprints, Cogprints).

This led to a list of 57 servers (Table 1). Included in our analysis are three preprint servers based on Open Science Framework (OSF) preprints infrastructure[38] that have during the study period announced leaving OSF (EarthArXiv, INA-Rxiv) or shutting down (MarXiv).[39]

Analysed variables

We analysed 7 topics related to preprint policies, 6 to submission requirements, and 18 to transparency in reporting and research integrity recommendations. Topic selection was informed by our previous work on journals' *Instructions to Authors*[33] and preprint metadata.[40] Mario Malički accessed webpages and stored their content as text files. By reading the web (sub)pages in full, he then indicated if the variable was addressed, and copied the text addressing a variable to Microsoft Excel. After this process was done for all servers, he then categorized how servers addressed these topics using an iterative process of open-coding and constant comparison. While initial text extraction was done at the time when the pages were accessed (time recorded in our raw database), extraction was double checked again from 15 April to 1 May 2020 using the stored text files. Two servers contained some pages in non-English language (ChinaXiv and ArabXiv) and those pages were Google translated. While SSRN is composed of 50 research networks, all point to the same instructions with only additional requirements for posting of clinical case reports. And so, we treated all SSRN networks as one server and included recommendations listed for case reports.

The variables extracted were:

Preprint policies:

1) Instruction to Authors: We checked if servers had a webpage that contained preprint submission information and extracted its title (e.g. *Instructions* to Authors, *Submission Guidelines*). We additionally checked if information on when the latest update to that page was made, and if a version number was assigned to the instructions.

2) Moderation: We extracted descriptions of screening or moderating procedures that preprints undergo before or after they are made available online.

3) Versioning: We extracted descriptions of updating or storing different versions of preprints.

4) Commenting: We checked if servers had platforms or forums that allow readers or other stakeholders to post comments about a preprint.

5) Preprint policies of journals: We checked if servers described or provided links to journalspolicies on posting preprints (e.g. *SHERPA/ROMEO* website,[21] Wikipedia's *List of academic journals by preprint policy*,[36] or servers' own compiled list).

6) Direct transfer: We checked if servers described the option to submit a preprint directly to a journal from the server, or vice versa.

7) Text mining: We checked if servers allowed text mining of deposited preprints.

Submission requirements:

1) Scope: We checked if servers listed scientific (sub)fields requirements.

2) Study type: We checked if specific article or study types were listed as (not) acceptable for deposit (e.g. reviews, case reports, opinion papers).

3) Preprint structure: We checked if a specific manuscript structure (e.g. IMRaD - Introduction, Methods, Results and Discussion) was recommended or required, or if any structure was implied (e.g. "*Papers should be laid out in the usual format of a scientific paper*").

4) Abstract guidelines: We checked if any recommendation or requirement for abstracts were provided, including for abstract's structure or length.

5) Reference style: We checked if any specific reference style was recommended or required.

6) (La)Tex submission: We checked if servers mentioned that preprints could be submitted as (La)Tex files in addition to PDFs, text or document files.

Transparency in reporting and research integrity recommendations:

1) Authorship: We checked if a definition or guidance on authorship was offered, including on the number of authors, their by-line order, or equal/shared contribution practice.

2) Conflicts of Interest: We checked if authors were asked to declare interest, irrespective of the terms used (e.g. conflicts of, competing or statement of interest).

3) COPE:We checked if *Committee on Publication Ethics (COPE)* was mentioned or recommended to authors regarding any aspects of publication ethics.

4) Data Sharing: We checked if data sharing (statement) was recommended or required.

5) Errata: We checked if corrections of preprints after deposit (i.e. errata or corrigenda) were mentioned, or if it was stated that any such changes should be made by posting a new version of a preprint.

6) Ethics Approval: We checked if authors were required to declare or provide proof on obtained ethics approval for their study, or that they conducted their study according to established standards (e.g. *Declaration of Helsinki)*.

7) Funding: We checked if authors were required to declare study funding, and if *Crossref Funder Registry* was recommended for correct nomenclature of funding bodies.[41]

8) ICMJE: We checked if *International Committee of Medical Journal Editors (ICMJE)* was referred to for any of their recommendations (e.g. manuscript formatting, trial registration, authorship, conflicts of interest, statistical guidance).

9) Image Manipulation: We checked if (screening for) image manipulation or duplication was described.

10) Limitations: We checked for recommendations or requirement of describing study limitations anywhere in the manuscript.

11) Null Results: We checked if servers stated that studies with null or negative results could be posted as preprints.

12) ORCID iD: We checked if an *Open Researcher and Contributor ID (ORCID iD)* was recommended or required for authors.

13) Patents: We checked if any special considerations were mentioned in case the content of the preprints should or was used for acquiring of patents.

14) Plagiarism: We checked if (screening for) plagiarism was mentioned and if the service or software used was specified.

15) Replication studies: We checked if any recommendations or restrictions were made for depositing replication studies.

16) Reporting guidelines: We checked if reporting guidelines were required or recommended. Additionally, we checked if the *Enhancing the QUAlity and Transparency Of health Research Network (EQUATOR)* was mentioned.

17) Statistical guidance: We checked whether any statistical reporting guidance was mentioned, including reporting of Bayes factors, confidence intervals, effect sizes, power or sample size calculations.

18) TOP guidelines: We checked whether *Transparency and Openness Promotion (TOP)* guidelines were mentioned.

Statistical analysis

We present the number and percentages (out of all analysed preprint servers) of servers addressing each extracted topic. We also present the median number of addressed topics across servers related to transparency in reporting and research integrity, as well as the accompanying interquartile range (IQR). Differences between the (median) number of addressed topics between servers that had *Instructions to Authors* webpage vs those without such page were compared using Mann-Whitney test.

Data Sharing

All extracted data is available on our project's data repository site: 10.17632/zrtfry5fsd.4.

Results

We analysed 57 preprint servers. While the majority of the servers (n=41, 72%) currently accept only specific (sub)disciplinary research, ten (18%) accept research from all disciplines, and six (11%) limit deposits to researchers coming from a specific region or a country (Table 1). Almost half of the servers (n=27, 47%) had a webpage (title) that could be categorized as *Instructions to Authors*. Preprints' policies and submission requirements that were most frequently described on servers' webpages and submission platforms were scope (n=57, 100%), moderation or

screening procedures servers employed before or after the preprints are made public (n=47, 82%), and advice for authors to check preprint policies of journals before submitting them for later publication (n=40, 70%). Of the 18 transparency in reporting and research integrity topics we analysed, preprints servers addressed a median of 1 topic (IQR 0 to 3). That number was slightly higher for servers with *Instructions to Authors* pages compared to those without (Md=2, IQR 2 to 6, vs Md=0, IQR 0 to 1, respectively, Mann Whitney P<0.001). Across all servers most commonly addressed topics were data sharing (n=22, 39%), plagiarism (n=15, 26%) and use of ORCID iD (n=14, 25%). Addressing of all topics across servers is presented in Table 2 and described below. Details per each server are available in our raw data file.

Detailed Results:

Preprint policies:

1) Instruction to Authors: Almost half of the servers (n=27, 47%) had a webpage (title) that could be classified as *Instructions to Authors* or *Submission Guidelines*. Two fifths of servers (n=23, 40%) covered some of the information traditionally covered in ItAs in other pages (e.g. *About* section, *FAQ*, *Policies*). Seven servers (all using the OSF preprint infrastructure), did not contain any pages except *Powered by OSF Preprints* link displayed below their name.

Four servers listed the date when their ItA was last updated, and only one, EarthArxiv, indicated an ItA version number (one more server, Mitofit, indicated a version number for their submission form template). One server, MediArXiv, had a note that their guidelines "*are subject to change*".

2) Moderation: Most servers (n=47, 82%) contained some information on their webpages or submission forms for screening/moderations checks they conduct before (n=39, 68%), or after the preprints are made public (n=8, 14%, all using the OSF preprint infrastructure). While over half (n=24, 51%) provided descriptions of checking for more than scope and scholarly nature of the preprints (e.g. checking for offensive langue or plagiarism), only two servers implied use of checklists, Preprints.org and Research Square, of which the latter displayed a pre-screen badge on every preprint webpage to indicate what was checked. Only one server, EcoEvoRxiv, indicated the number of individuals who perform checks for a specific preprint (one individual for research articles, and two for review or opinion articles). Preprints.org also had an invite and training for researchers willing to screen submissions.

3) Versioning: Little over half of servers (n=30, 53%) addressed the option to update or version preprints. Most servers did not impose limits on how many versions are allowed with Preprints-org stating encouraging authors to revise "*as often as they see fit*". RePec, on the other hand, advised authors to update preprints only for significant changes, and had a limit on first revision being allowed earliest 6 months after initial post. Lingbuzz allowed a grace period of one week to modify original submission, and arXiv stated that same date edits would not constitute new versions. ViXra stated that preprints could be replaced up to 9 times, but that a preprint is supposed to reach a "final form" (it also stated: "*If the number of replacements exceeds five then you are probably doing something wrong*"). Only OARR explicitly stated that "*All submission criteria apply to new version*". OARR also mentioned that any changes to supplementary materials also require posting a new version. Finally, arXiv recommends authors to indicate the nature of changes between versions, and that revision after version 5 will not be advertised to the community. SSRN only displays the newest version, except in cases when the history of previous versions is desired, or there is a difference in authors, language translation, or paper is included in multiple paper series.

4) Commenting: More than two thirds allowed commenting of preprints on their servers or forums (n=39, 68%), with Preprints.org also stating that it allows users to flag comments they find inappropriate, and MitoFit moderating comments before they are posted online.

5) Preprint policies of journals: More than two thirds of servers (n=40, 70%) advised authors to check journals' policies regarding preprints, most often referring them to the SHERPA/ROMEO website[21] or Wikipedia's *List of academic journals by preprint policy*.[36] ChemRxiv also indicated that *"majority of journals allow authors to first post preprints"*, while viXra more strongly indicated that "*no respectable journal*" would make publication conditional on the preprint's withdrawal, and PsyArXiv that preprints articles before publication (in a journal) "*are your creative product to do with as you please*"; with Preprints.org stating that "*preprints will not be removed to allow journal submission*" and that "*authors should check in advance whether the journal they intend to submit to accepts preprints.*"

6) Direct transfer: Ten servers (18%) had the ability to accept deposits from journals as well as forward preprints to journals for submission, with HAL also having the option to post deposits to arXiv.

7) Text mining: Seven servers (12%) stated they allow text mining of preprints, while one, e-LIS, prohibited it.

Submission requirements:

1) Scope: The majority of the servers (n=41, 72%) is meant for specific (sub)disciplinary research, 10 (18%) accept research from all disciplines, and 6 (11%) limit deposits to researchers coming from a specific region or a country (Table 1). Two servers addressed depositing of old documents, with e-LIS allowing deposits of those released in the public domain, and PhilSci of "*a small number of out of print texts in philosophy of science that are free of copyright encumbrance*". Only one server imposed the limit on the number of preprints that can be deposited (EdArXiv, with the limit of 30 preprints per calendar year), while arXiv stated that "*there is a practical limit to the rate at which appropriate, independent submissions can be produced by any one person*" and that they might limit posting for authors with excessive submission rates (though the rate was not specified).

2) Study type: More than half of servers (n=31, 54%) indicated the type of studies (not) accepted for deposit.

3) Preprint structure: A third of servers (n=19, 33%) indicated or implied preprint structure, with six (11%) indicating an IMRaD preprint structure (n=6, 11%), 11 (19%) implied a standard structure (e.g. ESSOar: "*equivalent to what is typically contained in a scholarly manuscript*"), with two (4%) servers recommending the style resembles that of the journal where authors plan to submit the preprint.

4) Abstract guidelines: A fifth of servers (n=12, 21%) recommended abstract structure or length, while one server, Preprints.org, also recommended the use of a graphical abstract alongside a text abstract.

5) Reference style: A reference style was recommended by eight (14%) servers, with an additional two servers allowing any reference style.

6) (La)Tex submission: All servers allowed preprints to be submitted as word or PDF documents, with about one fifth (n=10, 18%) allowing (La)Tex format.

Transparency in reporting and research integrity recommendations:

1) Authorship: Eight servers (14%) addressed authorship, four (7%) referring to ICMJE's guidelines for definition, two (4%) to those that deserve merit or made a substantial contribution, and one, ChinaXiv, advocated use of author contributions in preprint templates. One server, OSF Preprints, mentioned that authors should agree on the byline order. No servers addressed shared authorship.

One server, EconStore, specified that submitting authors "*must be employees of an academic institution (inside or outside a university)*", while another, SportRxiv, recommended including a Twitter handle on the title page "*to aid community interaction with the work*". Finally, ESSOar, specified that authors cannot be anonymous, and viXra that authors can use *"nom de plume"* if it is used for all research of that individual due to reasons like avoiding gender or ethnic bias, or avoiding employers or acquaintances judgement.

2) Conflicts of Interest: Nine servers (16%) required authors to declare conflicts, competing or statement of interest.

3) COPE:Two servers (4%) mentioned they would strive to follow all relevant guidance and best practices developed by Committee on Publication Ethics (COPE).

4) Data Sharing: Data sharing was recommended by 17 servers (30%), while an additional four (7%) required it.

5) Errata: Ten servers (18%) recommended that any changes to a preprint should be deposited as a new version of a preprint. One server, Lingbuzz, had a 7-day grace period for modification for new uploaded preprints (see more on versioning above), and e-Lis recommended that if author(s) cannot modify the copyright transfer agreement to allow self-archiving, they "*append or link a corrigenda file to the already self-archived preprint*".

6) Ethics Approval: Five servers (9%) required studies having an ethics approval, one, Preprints.org, that a study is also conducted according to the Declaration of Helsinki, and two additional servers required that ethics standards are followed.

7) Funding: Nine servers (16%) required declarations of funding. One server, OARR, recommended the use of Crossref Funder Registry for correct nomenclature of funding bodies.[41]

8) ICMJE: Five servers (9%) referred authors to International Committee of Medical Journal Editors (ICMJE) recommendations.

9) Image Manipulation: Two servers (4%) mentioned image manipulation as an inappropriate practice.

10) Limitations: Two servers (4%) had specific sections in their templates for reporting of study limitations.

11) Null Results: Six servers (11%) encouraged authors to deposit preprints that described null or negative results.

12) ORCID iD: Use of ORCID iD was recommended by 14 servers (25%).

13) Patents: Four servers (7%) mentioned patents, two stating that posting a preprint will compromise patent application, a second that any relation of a preprint to a patenting procedure must be disclosed, and a third that preprints can be used as proof of priority for patent applications.

14) Plagiarism: Plagiarism as inappropriate practice was mentioned by 15 servers (26%), eight of which (14%) stated that all preprints will be screened for plagiarism. Of the eight, only two specified the screening tool (both listing iThenticate).

15) Replication studies: Three servers (5%) encouraged preprinting of replication studies, one, EarthArXiv, only for software papers.

16) Reporting guidelines: Three servers (5%) recommended use of reporting guidelines, two advising authors to check EQUATOR Network, and one, SSRN, recommending use of CARE[42] for reporting of case reports.

17) Statistical guidance: No servers provided guidance for statistical reporting.

18) TOP guidelines: Only two servers referenced Transparency and Openness Promotion (TOP) guidelines on their server webpages (even though 26 servers use OSF preprints infrastructure we did not consider a link "*powered by OSF*" to be an endorsement of the TOP guidelines).

Discussion

Our analysis shows that even though most preprint servers employ screening checks for content that is posted on them, they provide very little guidance on issues that are important for transparency and research integrity. One possible reason behind this is the fact than less than half of the servers we analysed had webpages that resembled journals' *Instructions to Authors,* which in traditional scholarly publishing might contain such information. But even servers that had such webpages covered only a median of 2 out of 18 transparency in reporting and research integrity recommendations we analysed. Another reason could lie in main purpose of the preprint servers which is to provide a space for fast exchange of research and to, in the words of two servers in our sample, "*leave it to the individuals, communities, and institutions to develop their own criteria, announcements, journals, lists, and analyses of scholarly work*".[43, 44] And while not providing strict reporting or formatting requirements might save authors' time in preparing and uploading a preprint, it might also lead to decreased readability of preprints and the lack of details needed to assess their validity. The recent explosion of SARS-CoV-2 preprints and their social and media coverage has prompted several servers to address these issues by adding highly visible disclaimers which state that preprints should "*not be reported in news media as established information*" or be "*relied upon without context to guide clinical practice or health-related behaviour*".[45]

We have also shown that almost half of servers lacked detailed descriptions of what and how screening checks are conducted, including by how many (independent) individuals. Servers could benefit from clearly stating this information, and explore the use of stamps or badges (akin to ones used by *Research Square*) for each paper that would indicate passed screening checks. Additionally, several services recently emerged for listing and monitoring of journal policies on peer review, preprints, and research integrity,[21, 46, 47] as well as for automatic extraction and checks for transparency of reporting and adherence to integrity standards in published articles,[48-50] and it is therefore likely that the same will soon be applied to preprint servers, and individual preprints hosted on them. Clear guidance on reporting and good scholarly practices could also go a long way in helping raising awareness of those topics to young researchers, who, with the recent increase in the number of open science initiatives and the use of preprints, might end up posting a preprint before they ever submit a paper to traditional publishing venue. Producing standards and recommendations that would be agreeable to most scholars is not an easy task, and we are aware that some preprint servers are likely underfunded and depend on volunteers and willingness of the scholarly community to uphold them. Servers could therefore consider adapting the recently released universal template of *Instruction to Authors* to serve their purpose, and learn and share experiences with one another on reporting recommendations, as well as engage with their community, as some have done in the past when drafting their codes of conduct or exploring adding additional servers to servers like sharing of posters, presentations or other project outputs.[30, 51-53]

Our analysis has also shown that majority of servers also allow posting of public comments and in that way aim to encourage interaction among researchers and the public. Research on preprint discourse on social media and server commenting platforms is still scarce, but the recent exchange around SARS-CoV-2 preprints further exemplifies how fast peer review, retractions, but also conspiracy theories can emerge due to comments and interpretations of preprints in social media and press.[54] Finally, while we have identified 57 servers where researchers and deposit their

preprints, no resource yet exist that covers or aggregates their metadata or deposited full texts. We previously listed considerations for building such a resource, as well as recommendation for essential preprints metadata.[40]

Our research has several limitations. First, even though we read the full content of server webpages that could contain information on policies and instructions to authors, some practices and standards that are being applied are likely not listed there. Second, we chose to cover servers that did not limit posting of preprints based on funding or institutional affiliation or those that actively seek peer review, and policies of such servers might differ from the ones we covered here. Third, some of the webpage content was written in non-English language which we Google-translated, so we might have missed or misinterpreted some information on two servers (ChinaXiv and ArabiXiv). Fourth, we only explored some of the reporting and integrity topics, and additional topics warrant exploration in future studies, including how or whether analysed topics are addressed in individual preprints hosted on the servers.

In conclusion, with the recent increase in the number preprint servers and depositing of preprints, there is an opportunity for servers to encourage or require transparent reporting of research, adherence to research integrity standards, and stating of checks that are made before the preprints are made public. In doing so they could improve the quality and trust in the scholarly information exchange.

Literature

1. Moxham N, Fyfe A. The Royal Society and the prehistory of peer review, 1665–1965. The Historical Journal. 2017:1-27.
2. Spier R. The history of the peer-review process. TRENDS in Biotechnology. 2002;20(8):357-8.
3. Farrell PR, Magida Farrell L, Farrell MK. Ancient texts to PubMed: a brief history of the peer-review process. Journal of Perinatology. 2017;37(1):13-5.
4. Larson BP, Chung KC. A systematic review of peer review for scientific manuscripts. Hand. 2012;7(1):37-44.
5. Jefferson T, Rudin M, Brodney Folse S, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies. The Cochrane database of systematic reviews. 2007;2.
6. Bruce R, Chauvin A, Trinquart L, Ravaud P, Boutron I. Impact of interventions to improve the quality of peer review of biomedical journals: a systematic review and meta-analysis. BMC medicine. 2016;14(1):85.
7. Bornmann L, Mutz R, Daniel H-D. A reliability-generalization study of journal peer reviews: A multilevel meta-analysis of inter-rater reliability and its determinants. PloS one. 2010;5(12):e14331.
8. Himmelstein D. Satoshi Village: the blog of Daniel Himmelstein Internet2016. Available from: https://blog.dhimmel.com/history-of-delays/.
9. Björk B-C. Acceptance rates of scholarly peer-reviewed journals: A literature survey. El profesional de la información. 2019.
10. Maggio LA, Bynum WE, Schreiber-Gregory DN, Durning SJ, Artino AR. When will I get my paper back? A replication study of publication timelines for health professions education research. bioRxiv. 2019:783332.
11. Cobb M. The prehistory of biology preprints: A forgotten experiment from the 1960s. PLOS Biology. 2017;15(11):e2003995.
12. Li X, Thelwall M, Kousha K. The role of arXiv, RePEc, SSRN and PMC in formal scholarly communication. Aslib journal of information management. 2015;67(6):614-35.
13. Ginsparg P. ArXiv at 20. Nature. 2011;476(7359):145.

14. Chiarelli A, Johnson R, Richens E, Pinfield S. Accelerating scholarly communication: the transformative role of preprints. Zenodo. [Published Report]. In press 2019.

15. Rittman M. Preprint Servers. 2018. Available from: http://researchpreprints.com/.

16. Packer AL, Santos S, Meneghini R. SciELO in Perspective. 2017. Available from: https://blog.scielo.org/en/2017/02/22/scielo-preprints-on-the-way/.

17. News P. Public Knowledge Project 2020. Available from: https://pkp.sfu.ca/2020/02/24/the-road-to-preprints-part-1-introducing-open-preprint-systems/.

18. Therapeutics O. Therapoid 2020. Available from: https://therapoid.net/.

19. Lin J. Preprints growth rate ten times higher than journal articles: Crossref.org; 2018 updated September 11 2018. Available from: https://www.crossref.org/blog/preprints-growth-rate-ten-times-higher-than-journal-articles/.

20. Anaya J. Monthly Statistics 2018. Available from: http://www.prepubmed.org.

21. SHERPA RoMEO 2020. Available from: https://v2.sherpa.ac.uk/romeo/.

22. Wang Z, Glänzel W, Chen Y, editors. How self-archiving influences the citation impact of a paper. Proceedings of the 23rd International Conference on Science and Technology Indicators; 2018: CWTS, Leiden University; Leiden.

23. Serghiou S, Ioannidis JP. Altmetric Scores, Citations, and Publication of Studies Posted as Preprints. JAMA. 2018;319(4):402-4.

24. Fu DY, Hughey JJ. Releasing a preprint is associated with more attention and citations. bioRxiv. 2019:699652.

25. Haustein S, Bowman TD, Macaluso B, Sugimoto CR, Larivière V. Measuring Twitter activity of arXiv e-prints and published papers.

26. Fraser N, Momeni F, Mayr P, Peters I. The effect of bioRxiv preprints on citations and altmetrics. bioRxiv. 2019:673665.

27. Feldman S, Lo K, Ammar W. Citation Count Analysis for Papers with Preprints. arXiv preprint arXiv:180505238. 2018.

28. Klein M, Broadwell P, Farb SE, Grappone T. Comparing published scientific journal articles to their pre-print versions. International Journal on Digital Libraries. 2019;20(4):335-50.

29. Carneiro CFD, Queiroz VGS, Moulin TC, Carvalho CAM, Haas CB, Rayêe D, et al. Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. bioRxiv. 2020:581892.

30. Malički M, Jerončić A, ter Riet G, Bouter LM, Aalbersberg IJJ. Systematic Review of Publications Analysing Journals' Instructions to Authors. 2019.

31. Stevens A, Shamseer L, Weinstein E, Yazdi F, Turner L, Thielman J, et al. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. BMJ. 2014;348:g3804.

32. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review a. Systematic reviews. 2012;1(1):60.

33. Malicki M, Aalbersberg IJJ, Bouter L, Ter Riet G. Journals' instructions to authors: A cross-sectional study across scientific disciplines. PLOS One. 2019;14(9):e0222157.

34. Rittman M. Preprint Servers Internet2017. Available from: https://docs.google.com/spreadsheets/d/17RgfuQcGJHKSsSJwZZn0oiXAnimZu2sZsWp8Z6ZaYYo/edit#gid=0.

35. ASAPbio. Surveying the landscape of products and services for sharing preprints Internet2019. Available from: https://asapbio.org/preprint-products.

36. Wikipedia contributors. List of academic journals by preprint policy 2020. Available from: https://en.wikipedia.org/w/index.php?title=List_of_academic_journals_by_preprint_policy&oldid=940236037.

37. Stack Exchange community. Available from: https://academia.stackexchange.com/questions/84/preprint-services-other-than-arxiv-for-other-fields.

38. Science CfO. OSF Preprints accelerates scholarly review, publishing and discovery 2020. Available from: https://cos.io/our-products/osf-preprints/.

39. Mallapaty S. Popular preprint servers face closure because of money troubles. Nature. 2020;578(7795):349.

40. Malicki M, Alperin JP. Preprints Uptake and Use Project. [Blog]. In press 2019.

41. Crossref. Funder Registry. 2018. Available from: https://www.crossref.org/services/funder-registry/.

42. Gagnier JJ, Kienle G, Altman DG, Moher D, Sox H, Riley D, et al. The CARE guidelines: consensus-based clinical case reporting guideline development. Journal of medical case reports. 2013;7(1):223.

43. engrXiv Submission and Moderation Guidelines 2020. Available from: https://blog.engrxiv.org/guidelines/.

44. MediArXiv Moderation Policy 2020. Available from: https://mediarxiv.com/policies/.

45. COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv. bioRxiv. [Search Engine]. In press 2020.

46. Platform for Reponsibible Editorial Policies: Radboud University and Center for Science and Technology Studies. Available from: www.responsiblejournals.org.

47. New Measure Rates Quality of Research Journals' Policies to Promote Transparency and Reproducibility: Center for Open Science Available from: https://cos.io/about/news/new-measure-rates-quality-research-journals-policies-promote-transparency-and-reproducibility/.

48. UNSILO. UNSILO 2019. Available from: https://site.unsilo.com/site/unsilo-manuscript-evaluation/.

49. Hawrwood J. Penelope AI. Available from: https://www.penelope.ai/.

50. Serghiou S, Contopoulos-Ioannidis DG, Boyack KW, Riedel N, Wallach JD, Ioannidis JPA. Assessment of transparency indicators across the biomedical literature: how open is open? bioRxiv. 2020:2020.10.30.361618.

51. Ufnalska S, Terry A. Proposed universal framework for more user-friendly author instructions. European Science Editing. 2020;46:e53477.

52. Seeking Feedback on arXiv Code of Conduct 2019. Available from: https://arxiv.org/help/policies/code_of_conduct.

53. Rieger OY, Steinhart G, Cooper D. arXiv@ 25: Key findings of a user survey. arXiv preprint arXiv:160708212. 2016.

54. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of" infodemics" in response to COVID-19 epidemics. arXiv preprint arXiv:200403997. 2020.

Tables

Table 1. List of analysed preprint servers (N=57) and number of records they had on 29 May 2020.

All Disciplines	No. of records†	Subdisciplinary*	No. of records†
Authorea Preprint Repository	4,331	AgriXiv (Agriculture)	426
Hyper Articles en Ligne (HAL)	48,610	APSA Preprints (Political and Related sciences)	181
JMIR Preprints	7,888	BioHackrXiv (Biohackathons, Biogatherings, and Biocodefests)	5
OSF Preprints	17,174	BodoArXiv (Medieval studies)	60
Preprints.org (MDPI)	14,052	ChemRxiv (Chemistry)	4,872
Research Square	12,962	Cryptology ePrint Archive (Cryptology)	14,817
Social Science Research Network (SSRN)	802,602	Earth and Space Science Open Archive (ESSOAr) (Earth and Space Science)	936
Thesis Commons	1,009	EarthArXiv (Earth Science)	1,491
ViXra	35,827	EcoEvoRxiv (Ecology, Evolution and Conservation)	213
Zenodo	2,463	EconStor (Economics and Business Studies)	119,864
		ECSarXiv (Electrochemistry and Solid State Science and Technology)	123
Region/Country/Language specific	**No. of records†**	EdArXiv (Education)	470
AfricArXiv	127	Electronic Colloquium on Computational Complexity (Computational Complexity)	3,563
Arabixiv	432	e-LIS (Librarianship, Information Science and Technology)	941
ChinaXiv	13,682	engrXiv (Engineering)	1,026
FrenXiv	108	FocUS Arhive (Focused Ultrasound)	42
INA-Rxiv	16,641	LawArXiv (Law)	1,261
IndiaRxiv	112	LingBuzz (Linguistics)	5,113
		LIS Scholarship Archive (LISSA) (Library and Information Sciences)	265
(Multi)Disciplinary*	**No. of records†**	MarXiv (Marine and Climate Studies)	449
Advance (SAGEpub)	544	Mathematical Physics Preprint Archive (Mathematical Physics and Related areas)	9,601
arXiv	1,708,255	MediArXiv (Media, Film, and Communication Studies)	91
bioRxiv	84,009	MetaArXiv (Meta-Research)	131
Commons Open Repository Exchange (CORE)	451	MindRxiv (Mind and Contemplative practices)	197
medRxiv	5,935	MitoFit Preprint Archives (Mitochondrial Physiology and Bioenergetics and Ergodynamics)	16
SocArXiv	5,497	NutriXiv (Nutritional Sciences)	69
SSOAR	8,193	OARR: Open Anthropology Research Repository (Anthropology)	58
		Optimization Online (Optimization and Related Topics)	7,531
		PaleorXiv (Paleontology and Paleobiology)	166
		PhilArchive (Philosophy)	48,927
		PhilSci-Archive (Philosophy of science)	4,210
		PsyArXiv (Psychology)	9,475
		RePEc / Munich Personal RePEc Archive (Economy)	49,164
		SportRxiv (Sport, Exercise, Performance, and Health research)	200

*We use the following main disciplines categorization: Arts & Humanities, Health Sciences, Life Sciences, Physical Sciences, and Social Sciences.

† We intentionally use the term number of records as not all servers have filters that clearly differentiate between preprints and published papers or account for duplicate records.

Table 2. Preprint servers' policies, submission requirements and transparency in reporting and research integrity reccomendations (N=57).

Preprint Policies	n	%
Instruction to Authors	27	47%
Moderation	46	81%
before preprint is made public	38	67%
after preprint is made public	8	14%
Versioning	30	53%
Commenting	39	70%
Preprint policies of journals	40	70%
SHERPA/ROMEO	36	63%
Wikipedia List	12	21%
other	4	7%
Direct transfer	10	18%
from a preprint server to a journal	7	12%
from a journal to a preprint server	2	4%
both directions	1	2%
Text mining	8	14%
allowed	7	12%
prohibited	1	2%

Submission requirements	n	%
Scope	57	100%
(sub)disciplinary	41	72%
all disciplines	10	18%
region/country specific	6	11%
Study Type	31	54%
Preprint structure	19	33%
implied structure (e.g. usual scientific format)	11	19%
IMRaD	6	11%
style of a journal where it will be published	2	4%
Abstract guidelines	12	21%
minimal or maximal word limit	10	18%
structured abstract	3	5%
recommend use of a graphical abstract	1	2%
Reference style	16	28%
specified style	8	14%
no style specified but preprint should contain ref.	6	11%
any style	2	4%
(La)Tex submission	10	18%

Transparency in reporting and research integrity reccomendations	n	%
Authorship	8	14%
ICMJE definition	2	4%
those who merit / made substantive contribution	2	4%
when no disciplinary criteria exist use ICMJE's	1	2%
ICMJE and CREDIT taxonomy recommended	1	2%
specify authors' contributions	1	2%
determine the order in which the names should appear	1	2%
Conflicts of Interest	9	16%
Crossref Funder Registry for funder names	1	2%
Committee of Publication Ethics (COPE)	2	4%
Data Sharing	22	39%
recommended	17	30%
required	4	7%
allow linking or uploading data	1	2%
Errata	12	21%
Changes should be submitted as new version	10	18%
One week grace period to make changes	1	2%
Errata can be deposited for published papers	1	2%
Ethics Approval	9	16%
required	5	9%
must follow the Declaration of Helsinki	1	2%
patient consent must be addressed for case reports	1	2%
abide by appropriate ethical standards	2	4%
Funding	9	16%
ICMJE	5	9%
Image Manipulation	2	4%
Limitations	2	4%
Null Results	6	11%
ORCID iD	14	25%
Patents	3	5%
preprint will compromise your patent application	1	2%
preprint related to a patent (application) must be disclosed	1	2%
preprints can be used as proof of priority for patents	1	2%
Plagiarism	15	26%
screening using an unspecified tool	6	11%
screening using iThenticate	2	4%
Replication studies	3	5%
(largely) replicate or reproduce previous work	2	4%
replication studies for software papers	1	2%
Reporting guidelines	3	5%
check Equator Network	2	4%
use CARE for case reports	1	2%
Statistical guidance	0	0%
TOP guidelines	2	4%

Declarations

Changes between this preprint and the published version: The manuscript was shortened and formatted to JAMA's research letter requirements. During the peer review process, information on number of health sciences preprints hosted on the servers, as well as subgroup analyses for servers hosting more than 500 health sciences preprints were added. Those health servers addressed a median of 5 out of 18 transparency recommendations, which was higher than a median of 1 found for all other servers (for details see the published version).

Statement of Interests: IJsbrand Jan Aalbersberg is Senior Vice-President of Research Integrity for Elsevier, and Elsevier owns the preprint server SSRN. Gerben ter Riet received research grants from Elsevier.

Ethics Approval: No ethics approval was needed as we analysed publicly available data.

Data Availability: All data and project notes are available at our project's data repository site: 10.17632/zrtfry5fsd.4