

# Duplicated or cloned preprints: the lack of across-the-board ethical regulation of preprint servers may risk leading to citation-abusive and *curriculum vitae*-enhancing practices

Jaime A. Teixeira da Silva  
Independent researcher, Miki-cho, Japan  
jaimetex@yahoo.com  
ORCID 0000-0003-3299-2772

## Abstract

Preprints have become an attractive form of publication in recent years since they allow academics to advance the speed of publication by being able to publish almost instantaneously an often crude form of work, or research in progress, while offering a quantitative substitute to showcase early research efforts. Despite this positive set of aspects, outside of the intense debate of the role that preprints played in COVID-19 research, very little is being discussed in academic circles whether this purpose and privilege can be abused. This paper tempers over-enthusiastic preprint proponents by cautioning researchers who may seek to overly amplify the visibility of their research efforts using preprints, specifically by abusing multiple preprint servers to clone the same work as many times as possible, taking advantage of the absence of stringent preprint ethics, moderation and verification policies, to gain greater exposure, readership and citations of their work. At the heart of this exercise lies the question: What is the limit to the number of copies of work, disseminated via preprints, that is allowed or acceptable? Several cases are provided to illustrate these arguments and risks.

**Keywords:** ethics, insufficient regulation, peer-reviewed literature, policies, transparency

## Why have preprints become such an attractive, but risky, publication format?

Even though preprints are frequently touted as a rapid way to disseminate information and to gain open and public input into non-peer-reviewed work, this bold objective has not always fully materialized in the biomedical sciences, leaving a body of biomedical researchers skeptical about the validity of this fairly unregulated form of publication<sup>1,2</sup>. There are some technical limitations to preprints, such as persistence and discoverability, while social acceptance has allowed them to be more practically useful in some fields of research than in others, causing skewed use<sup>3</sup>. Incidentally, the second author of that latter paper leads a preprint advocacy organization, ASAPbio.

<sup>1</sup> Annette Flanagan, Phil B. Fontanarosa and Howard Bauchner, "Preprints involving medical research – do the benefits outweigh the challenges?", *JAMA* 324 (18) (2020): 1840–1843, doi: 10.1001/jama.2020.20674.

<sup>2</sup> Fernando Alfonso and Filippo Crea, "Preprints: a game changer in scientific publications?", *European Heart Journal* 44 (3) (2023): 171–173, doi: 10.1093/eurheartj/ehac665.

<sup>3</sup> Naomi C. Penfold and Jessica K. Polka, "Technical and social issues influencing the adoption of preprints in the life sciences", *PLoS Genetics* 16 (4) (2020): e1008565, doi: 10.1371/journal.pgen.1008565.

Minimalist screening by moderators, managers or other individuals responsible for managing the quality control of preprints submitted to their servers has resulted in an infusion of preprints plagued with ethical issues, such as plagiarism, duplication, academic fraud, and misconduct<sup>4</sup>, no different from the same issues observed in peer-reviewed papers. In fact, in 2023, only 8% of 36 ASAPbio-listed preprint servers had a page dedicated to or containing comprehensive ethics policies<sup>5</sup>. Evidently, absent ethical stringency and consistent across-the-board ethics policies, this invites abuse, and absent ethical repercussions for those abuses, it is only natural to expect that abusive practices will rapidly proliferate, making preprint servers potentially unreliable sources of scientific information, as was assessed in a survey, thus damaging their credibility as a publication format<sup>6</sup>.

Evidence of this abuse is not in short supply. The ease of publication allowed by preprint servers, their lax screening and often vague ethical policies, combined with a sense of urgency (or enthusiasm), especially among biomedical researchers who explored the reasons for and solutions to the COVID-19 pandemic, invited a wealth of unreliable preprints to be published, a solid portion of those eventually being retracted<sup>7</sup>. Another serious phenomenon that is not being intensively studied, but that is of essence to library and information scientists who cherish the permanence of the academic and scholarly record, is the opaque "disappearance" of metadata, background information and sometimes even the entire preprint itself, without any transparent explanation, and despite carrying a digital object identifier (DOI), despite the use of this digital identifier in ensuring the permanence of published literature. A term, the "silent retractions" (partial or full) of preprints, was coined for this phenomenon<sup>8</sup>, with specific case studies and respective deliberations having been recorded for Elsevier's *SSRN*<sup>9</sup> and SAGE's *Advance*<sup>10</sup>, two popular preprint servers. Much more attention is needed to this phenomenon, which is a clearly disruptive technology to library and information science, and is thus antithetic to its core principle of information preservation.

In other words, ethical principles underlying the use of preprints still remain highly unregulated and uncoordinated, including among status quo publishers that allow their use by authors who submit papers to their peer-reviewed journals, including as preprint server-publisher pairs to, as a possible hypothesis, facilitate the flow of information and knowledge to their own preprints, e.g., Elsevier + *SSRN*, Springer Nature + *Research Square*, SAGE + *Advance*, Wiley + *Authorea*, MDPI + *Preprints.org*, etc. In order to prevent the abuse of preprints to over-amplify one's own work, which might cause a negative image among peers, it was argued that preprints and peer-reviewed papers need to be treated as "ethical equals", i.e., ethical infractions that typically result in penalties for the latter, such as retractions, need to be equally applied to the

<sup>4</sup> Mario Malički et al., "Preprint servers' policies, submission requirements, and transparency in reporting and research integrity recommendations", *JAMA* 324 (18) (2020): 1901–1903, doi: 10.1001/jama.2020.17195.

<sup>5</sup> Jaime A. Teixeira da Silva and Serhii Nazarovets, "Current ethics policies in 36 preprint servers: Relevance for Academic Medicine", *Academic Medicine* 99 (2) (2024): 129–130, doi: 10.1097/ACM.0000000000005507.

<sup>6</sup> Courtney K. Soderberg, Timothy K. Errington and Brian A. Nosek, "Credibility of preprints: An interdisciplinary survey of researchers", *Royal Society Open Science* 7 (10) (2020): 201520, doi: 10.1098/rsos.201520.

<sup>7</sup> Joan E. Dodgson, "The pandemic has brought too much change: Too many preprints; too many retractions", *Journal of Human Lactation* 38(2) (2022): 207–208, doi: 10.1177/08903344221082571.

<sup>8</sup> Jaime A. Teixeira da Silva, "Silently withdrawn or retracted preprints related to Covid-19 are a scholarly threat and a potential public health risk: Theoretical arguments and suggested recommendations", *Online Information Review* 45 (4) (2021): 751–757, doi: 10.1108/OIR-08-2020-0371.

<sup>9</sup> Jaime A. Teixeira da Silva, "The 'silent' removal of bibliometric information of three *SSRN* preprints related to peer review, and then their full reinstatement", *Preservation, Digital Technology & Culture* 52(3) (2023): 85–89, doi: 10.1515/pdct-2023-0021.

<sup>10</sup> Jaime A. Teixeira da Silva, "An attempt to explain the 'silent' withdrawal or retraction of a SAGE *Advance* preprint", *Publishing Research* 2 (2023): 4, doi: 10.48130/PR-2023-0004.

former<sup>11</sup>. Absent a strict and equalized stratification of ethics for preprints and peer-reviewed papers, the former will surely be abused if abusive rule non-abiding authors perceive an ethical loophole that allows them to act unethically, and with impunity.

At the heart of this paper, and what some preprint proponents fail to recognize (or perhaps conveniently do not wish to recognize) is that preprint servers are currently ethically under-regulated. In this paper, I thus wish to temper over-enthusiasm for this form of publication by cautioning researchers who may seek to overly amplify the visibility of their research efforts using preprints, as a way to boost publication counts and quantitatively enhance their *curriculum vitae* or demonstrate that they are highly productive.

### **How many copies of a preprint can or should be published? Where does the limit lie?**

In the world of peer-reviewed science, a seemingly arbitrary rule has been imposed on academia that limits their ability to disseminate their work and thus knowledge through only a single published copy, and thus potentially restricting access to it, as well as dissemination of information in it. An exception are papers published in different languages (i.e., translated versions) or duplicated copies for which editors have granted special republication rights. Thus, except for exceptional cases, academics who might duplicate their work across two or more journals or venues are often labelled as unethical, and duplicated copies – if detected – face retraction, a decision that seems to be related as much to the issue of copyright infringement as an ethics infraction (i.e., duplicate submission and publication, or redundant publication). The premise of this ethical ruling is rarely challenged, and is generally imposed on academia, without any challenge by academics. Despite this, seemingly powerful groups, lobbyists, self-proclaimed ethicists and policy-makers are immune to this one-copy rule, in which any number of copies of a work can be published. In such cases, there is an ethical caveat, namely that provided that there is agreement among the “group”, this will ensure that multiple copies of, as an example, “guidelines” (which are still in essence academic papers published in peer-reviewed journals, not unlike the peer-reviewed papers by “regular” authors) can be published, often accompanied by a footnote indicating that multiple copies exist. One prominent example are the hepta-copies of the PRISMA guidelines<sup>12</sup>. This critical background is extremely important, because if exceptions to existing ethical rules are already occurring in standard academic publishing, then it can be expected that even worse abuses (or exceptions) may be observed in the highly unregulated (or poorly regulated) world of preprints.

Thus, at the heart of this mental exercise in this paper lies the question: What is the limit to the number of copies of work, disseminated via preprints, that is allowed or acceptable? For those that believe that such practices are not abusive, i.e., that multiple copies of preprints are allowed, but that multiple copies of peer-reviewed work are not allowed, thereby creating a dual set of ethical standards in academic publishing, then where does one draw the line for preprints? What should be the limit to the number of copies be that an academic is allowed to clone across preprint servers: 1, 2, 5, 10, 36, 100? Moreover, who has the right and the ethical prowess to make that ruling and determination on behalf of the global academic community? The answers to these questions might only arise after open debate among academics, and once stringent across-the-board ethical policies have been implemented by preprint servers. Even though ASAPbio, which wields considerable influence in the world of preprints, could make a decisive push to ensure that all of the preprint servers it indexes, are forced to implement strict

<sup>11</sup> Jaime A. Teixeira da Silva, “Should preprints and peer-reviewed papers be assigned equal status?”, *Journal of Visceral Surgery* 159 (5) (2022): 444–445, doi: 10.1016/j.jviscsurg.2022.08.003.

<sup>12</sup> Jaime A. Teixeira da Silva and Timothy Daly, “Against over-reliance on PRISMA guidelines for meta-analytical studies”, *Rambam Maimonides Medical Journal* 15 (1) (2024): e0004, doi: 10.5041/RMMJ.10518.

ethics policies, curiously, limited efforts have seemingly been made. That lack of effort to ensure stringent ethics policies in preprints in itself is worthy of an epistemic debate. Even so, as so often occurs with ethical policies and guidelines, managers of preprint servers are advised to liaise with academia and grass-roots academics to reach consensus prior to imposing new rules and regulations.

### Select case studies

Preprints often carry philanthropic or political motivations, and thus may be subjected to abuses or biases<sup>13</sup>, so it is important to debate case studies such as those in this section of this paper to appreciate why the absence of stringent ethics policies, even though preprint proponents like ASAPbio are also proponents of open science, and thus strong defenders of strict ethical policies, also cases of “ethical irony” in science publishing. The examples provided in Tables 1 and 2 (status last verified on 25 March 2025) do not represent a thorough or comprehensive list of examples, nor was a systematic methodology used to detect or record them. The sole purpose of listing these examples is to show that they exist, to indicate the number of copies that have been detected, to open up this finding to the academic community for wider debate, to allow relevant ethicists and policy makers to appreciate the delinquent nature of ethics policies that currently exist in a wide selection of preprint servers, and to appreciate the failure to standardize them across such servers. It is important that the authors of these multi-copy preprints also be given an opportunity to explain their rationale for cloning their preprints across multiple preprint servers, while the managers and owners of preprint servers also need to defend the apparent absence of ethics policies, or lax regulation and verification regarding this issue. Absent the input of these two parties or stakeholders, a fruitful debate will not result, and preprint servers may be used to amplify and disseminate publishing activities that may be deemed or perceived as unethical.

Table 1 indicates six cases of two copies of each preprint, whereas Table 2 makes note of a more exquisite case involving seven preprint servers. In all cases, the purpose appears to be to over-amplify the visibility of those papers, and thus, by association, the authors' *curriculum vitae*. Not much more can – or should – be extrapolated beyond these observations. Most importantly, these cases are not an accusation of unethical behavior in the sense that if ethics policies in preprint servers are non-existent or unclear, then authors who may select to publish the same paper across multiple preprint servers might not necessarily be doing so in breach of any stated ethical codes of conduct. However, are such practices fair, and are such publishing actions morally defensible? For example, whereas a single copy of a preprint might allow for a defined number of readers and citations, how would multiple copies of a work across two or more preprint servers amplify visibility and thus readership and citations of that cloned preprint?

<sup>13</sup> Jaime A. Teixeira da Silva, “The preprint wars”, *AME Medical Journal 2* (2017): 74, doi: 10.21037/amj.2017.05.23.

Table 1 List of six cases of preprints with two copies

Case #	DOI-based URLs of preprint (or post-print) copies	Preprint servers (owner/publisher)
1	Copy 1: <a href="https://doi.org/10.1101/2020.02.27.967588">https://doi.org/10.1101/2020.02.27.967588</a> Copy 2: <a href="https://doi.org/10.1101/2020.02.24.20025437">https://doi.org/10.1101/2020.02.24.20025437</a>	bioRxiv (Cold Spring Harbor Laboratory) <sup>1</sup> medRxiv (Cold Spring Harbor Laboratory) <sup>2</sup>
2	Copy 1: <a href="https://doi.org/10.1101/2020.11.04.367896">https://doi.org/10.1101/2020.11.04.367896</a> Copy 2: <a href="https://doi.org/10.1101/2020.10.30.20223180">https://doi.org/10.1101/2020.10.30.20223180</a>	bioRxiv (Cold Spring Harbor Laboratory) <sup>1</sup> medRxiv (Cold Spring Harbor Laboratory) <sup>2</sup>
3	Copy 1: <a href="https://doi.org/10.1101/2020.10.10.20210666">https://doi.org/10.1101/2020.10.10.20210666</a> Copy 2: <a href="https://doi.org/10.21203/rs.3.rs-125016/v1">https://doi.org/10.21203/rs.3.rs-125016/v1</a>	medRxiv (Cold Spring Harbor Laboratory) <sup>2</sup> Research Square (Springer Nature)
4	Copy 1: <a href="http://dx.doi.org/10.2139/ssrn.3849393">http://dx.doi.org/10.2139/ssrn.3849393</a> Copy 2: <a href="https://doi.org/10.31124/advance.14610291.v1">https://doi.org/10.31124/advance.14610291.v1</a>	SSRN (Elsevier) Advance (SAGE)
5	Copy 1: <a href="https://doi.org/10.48550/arXiv.2304.11215">https://doi.org/10.48550/arXiv.2304.11215</a> Copy 2: <a href="http://dx.doi.org/10.2139/ssrn.4423874">http://dx.doi.org/10.2139/ssrn.4423874</a>	arXiv (Cornell University) SSRN (Elsevier)
6	Copy 1: <a href="https://doi.org/10.21203/rs.3.rs-266574/v1">https://doi.org/10.21203/rs.3.rs-266574/v1</a> Copy 2: <a href="https://doi.org/10.1101/2021.01.27.21250617">https://doi.org/10.1101/2021.01.27.21250617</a>	Research Square (Springer Nature) medRxiv (Cold Spring Harbor Laboratory) <sup>2</sup>

<sup>1</sup> "Posting to both *bioRxiv* and *medRxiv* is not permitted and will result in article withdrawal. Manuscripts already present on other preprint servers will not be posted." <https://www.biorxiv.org/about/FAQ> (despite this ethics clause, the duplicate preprint stands, and the duplicate has not been withdrawn/retracted)

<sup>2</sup> "Posting to both *medRxiv* and *bioRxiv* is not permitted and will result in article withdrawal. Manuscripts already present on other preprint servers will not be posted." <https://www.medrxiv.org/about/FAQ> (despite this ethics clause, the duplicate preprint stands, and the duplicate has not been withdrawn/retracted)

Table 2 List of preprint servers on which one paper<sup>1</sup> has been archived as a preprint or as a post-print

DOI (or URL where no DOI exists)	Preprint server (owner)
<a href="https://doi.org/10.20944/preprints202307.2060.v1">https://doi.org/10.20944/preprints202307.2060.v1</a>	Preprints.org (MDPI)
<a href="https://doi.org/10.14293/PR2199.000283.v1">https://doi.org/10.14293/PR2199.000283.v1</a>	ScienceOpen Preprints (Science Open)
<a href="https://doi.org/10.48550/arXiv.2307.14810">https://doi.org/10.48550/arXiv.2307.14810</a> <sup>2</sup>	arXiv (Cornell University)
<a href="http://dx.doi.org/10.2139/ssrn.4530612">http://dx.doi.org/10.2139/ssrn.4530612</a>	SSRN (Elsevier)
<a href="https://doi.org/10.48550/arXiv.2307.14810">https://doi.org/10.48550/arXiv.2307.14810</a> <sup>2</sup>	OSF preprints (Center for Open Science)
<a href="https://easychair.org/publications/preprint/FbQT">https://easychair.org/publications/preprint/FbQT</a> <sup>3</sup>	Easy Chair (easychair.org)
<a href="https://www.researchgate.net/publication/373683915">https://www.researchgate.net/publication/373683915</a> <sup>3</sup>	ResearchGate preprint (ResearchGate)

<sup>1</sup> Matthew James Stephenson, "A differential datalog interpreter", *Software* 2 (3) (2023): 427–446. doi: 10.3390/software2030020 (with an expression of concern: doi: 10.3390/software3020011).

<sup>2</sup> Retracted/withdrawn. The retraction note only indicates that: "This submission has been removed due to violation of *arXiv* author policy".

<sup>3</sup> Apparent archival of now-retracted /withdrawn *arXiv* preprint.



## Conclusion

Preprints are touted as a way to reduce publication bias<sup>14</sup>, but the cases highlighted in Tables 1 and 2 point towards the potential for amplifying publication bias and the potential to also abuse preprint servers to clone copies of work in unlimited numbers due to the absence of across-the-board ethical guidelines, regulation, and verification. In Table 2, this abuse even extends itself to a key player of the open science movement, *OSF preprints*, hosted by the Center for Open Science (COS). Of the 29 preprint servers hosted by COS in 2024, about half had either ceased to be hosted on the COS platform, or had ceased publishing altogether<sup>15</sup>. Even where ethics guidelines exist that "outlaw" the cloning or duplication of preprints (e.g., *bioRxiv* and *medRxiv*), absent retraction, these rules exist merely on paper because they are not practically implemented, as noted in three cases in Table 1. Even if a researcher has the ability to post unlimited copies of the same work across two or multiple preprint servers, should they? In the author's opinion, even though such a possibility is attractive (after all, which academic does not wish to have more visibility, reads, and citations?), is it the correct mind-set for academic and scholarly publishing?

## References:

1. Flanagan, Annette, Phil B. Fontanarosa and Howard Bauchner. "Preprints involving medical research – do the benefits outweigh the challenges?". *JAMA* 324 (18) (2020): 1840–1843. doi: 10.1001/jama.2020.20674.
2. Alfonso, Fernando and Filippo Crea. "Preprints: a game changer in scientific publications?". *European Heart Journal* 44 (3) (2023): 171–173. doi: 10.1093/eurheartj/ehac665.
3. Penfold, Naomi C. and Jessica K. Polka. "Technical and social issues influencing the adoption of preprints in the life sciences". *PLoS Genetics* 16 (4) (2020): e1008565. doi: 10.1371/journal.pgen.1008565.
4. Malički, Mario et al. "Preprint servers' policies, submission requirements, and transparency in reporting and research integrity recommendations". *JAMA* 324 (18) (2020): 1901–1903. doi: 10.1001/jama.2020.17195.
5. Teixeira da Silva, Jaime A. and Serhii Nazarovets. "Current ethics policies in 36 preprint servers: Relevance for Academic Medicine". *Academic Medicine* 99 (2) (2024): 129–130. doi: 10.1097/ACM.0000000000005507.
6. Soderberg, Courtney K., Timothy K. Errington and Brian A. Nosek. "Credibility of preprints: An interdisciplinary survey of researchers". *Royal Society Open Science* 7 (10) (2020): 201520. doi: 10.1098/rsos.201520.
7. Dodgson, Joan E. "The pandemic has brought too much change: Too many preprints; too many retractions". *Journal of Human Lactation* 38 (2) (2022): 207–208. doi: 10.1177/08903344221082571.
8. Teixeira da Silva, Jaime A. "Silently withdrawn or retracted preprints related to Covid-19 are a scholarly threat and a potential public health risk: Theoretical arguments and suggested recommendations". *Online Information Review* 45 (4) (2021): 751–757. doi: 10.1108/OIR-08-2020-0371.
9. Teixeira da Silva, Jaime A. "The 'silent' removal of bibliometric information of three SSRN preprints related to peer review, and then their full reinstatement". *Preservation, Digital Technology & Culture* 52 (3) (2023): 85–89. doi: 10.1515/pdte-2023-0021.

<sup>14</sup> Kyu Jin Chung, "Preprints: What is their role in medical journals?", *Archives of Plastic Surgery* 47 (2) (2020): 115–117, doi: 10.5999/aps.2020.00262.

<sup>15</sup> Jaime A. Teixeira da Silva, "A perspective on the Center for Open Science (COS) preprint servers", *Science Editor and Publisher* 9 (1) (2024): 86–95, doi: 10.24069/SEP-24-05.

10. Teixeira da Silva, Jaime A. "An attempt to explain the 'silent' withdrawal or retraction of a SAGE Advance preprint". *Publishing Research* 2 (2023): 4. doi: 10.48130/PR-2023-0004.
11. Teixeira da Silva, Jaime A. "Should preprints and peer-reviewed papers be assigned equal status?". *Journal of Visceral Surgery* 159 (5) (2022): 444–445. doi: 10.1016/j.jvisc Surg.2022.08.003.
12. Teixeira da Silva, Jaime A. and Timothy Daly. "Against over-reliance on PRISMA guidelines for meta-analytical studies". *Rambam Maimonides Medical Journal* 15 (1) (2024): e0004. doi: 10.5041/RMMJ.10518.
13. Teixeira da Silva, Jaime A. "The preprint wars". *AME Medical Journal* 2 (2017): 74. doi: 10.21037/amj.2017.05.23.
14. Chung, Kyu Jin. "Preprints: What is their role in medical journals?". *Archives of Plastic Surgery* 47 (2) (2020): 115–117. doi: 10.5999/aps.2020.00262.
15. Teixeira da Silva, Jaime A. "A perspective on the Center for Open Science (COS) preprint servers". *Science Editor and Publisher* 9(1) (2024): 86–95. doi: 10.24069/SEP-24-05.

## Duplirane ili klonirane predpublikacije: manjak opštevažećih etičkih pravila za servere sa predpublikacijama može dovesti do zloupotreba citiranja i veštačkog poboljšavanja profesionalnih biografija

### Sažetak

Predpublikacije su u novije vreme postale privlačan žanr jer članovima akademske zajednice omogućavaju da ubrzaju objavljivanje tako što gotovo trenutno obelodanjuju često sirovu verziju svog rada ili istraživanja koje je u toku, nudeći istovremeno kvantitativnu zamenu za izlaganje istraživačkih aktivnosti u ranom stadijumu. Osim pozitivnih strana, i ostavljajući po strani intenzivnu raspravu o ulozi koju su predpublikacije imale u istraživanjima virusa COVID-19, u akademskim krugovima se malo razgovaralo o tome da li se njihova svrha i prednosti mogu zloupotrebiti. Ovaj rad ublažava preterani entuzijazam zagovornika predpublikacija tako što upozorava da će se možda pojaviti istraživači koji će pokušati da njihovim korišćenjem neopravdano povećaju vidljivost svojih istraživanja, naročito tako što će zloupotrebiti više servera za predpublikacije klonirajući isti rad što je više puta moguće i koristiti odsustvo jasne etike postupanja sa njima, kao i usvojene procedure za moderisanje i verifikovanje, da bi povećali izloženost, publiku i broj citiranja svog rada. U središtu ovog oglada stoji pitanje: Koliko je najviše kopija rada dozvoljeno ili prihvatljivo da se diseminiraju kao predpublikacije? Predstavljeno je više slučajeva koji ilustruju ovu argumentaciju i opasnosti u vezi sa serverima za predpublikacije.

**Ključne reči:** etika, nedovoljna regulacija, recenzirana literature, usvojene procedure, transparentnost

Примљено: 1. октобра 2024.  
Исправке рукописа: 23. марта 2025.  
Прихваћено за објављивање: 2. априла 2025.